

Dannemann, Bernhard C.

Conference Paper

Peer effects in secondary education: Evidence from trends in mathematics and science study 2015 based on weak-tie bonds

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2019: 30 Jahre Mauerfall - Demokratie und Marktwirtschaft - Session: Education Economics II, No. B17-V1

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Dannemann, Bernhard C. (2019) : Peer effects in secondary education: Evidence from trends in mathematics and science study 2015 based on weak-tie bonds, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2019: 30 Jahre Mauerfall - Demokratie und Marktwirtschaft - Session: Education Economics II, No. B17-V1, ZBW - Leibniz-Informationszentrum Wirtschaft, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/203485>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

PEER EFFECTS IN SECONDARY EDUCATION:
EVIDENCE FROM TRENDS IN MATHEMATICS AND SCIENCE STUDY 2015
BASED ON WEAK-TIE BONDS *

Bernhard C. Dannemann[†]

THIS VERSION: FEBRUARY 2019

Abstract

Peer effects in education production functions are predominantly employed as the mean classroom performance. Based on sociological theory and using spatial regression techniques, I introduce social network matrices that correspond to a weighting scheme for peers within the class. In a spatial regression, I show the presence of peer effects for the 8th grade population in the USA in the TIMSS 2015 student assessment. For students, the likelihood of cooperation increases conditional on visible and non-visible characteristics, i.e., age, gender, foreign origin, or scholastic achievement. Externalities of scholastic performance show a larger magnitude than the direct effects on the individual. The results are robust to various definitions of how peer groups are created, e.g., classroom-average, status homophily or value homophily and to the inclusion of school level fixed effects.

Keywords: Human Capital, Cognitive Skills, Peer Effects, Spatial Model,
Class Heterogeneity, Education Production Function

JEL Classification Numbers: I21, D62, C31, C18, D91

*I would like to thank my supervisor, Jürgen Bitzer, for helpful comments and suggestions. All remaining errors are my own.

[†]Carl von Ossietzky University Oldenburg, School of Computing Science, Business Administration, Economics, and Law (Faculty II), Institute of Economics, Building A5, 26111 Oldenburg, Germany, Tel.: +49-441-798-2625, website: <http://www.uni-oldenburg.de/en/vwl/applied-macroeconomics/team/bernhard-dannemann/>, e-mail: bernhard.dannemann@uni-oldenburg.de

1 Introduction

The role of peers in education is undoubtedly of importance as school environments regarding socio-economic composition have been accepted as an important factor in the acquisition and formation of cognitive skills in education in and out of school. While the usual determinants of individual student achievement are well researched, the way students cooperate and interact is frequently disregarded (Hanushek and Woessmann, 2011). These so-called peer effects are often - if considered - implemented as averages of peer performance, without an appropriate differentiation by class or real peer groups, but aimed at school or grade level, instead (Angrist, 2014; Lin, 2010). Furthermore, a large share of the discussion on peer effects has focused on technical problems of identification and estimation. However, clarifying the modes of effect and transmission channels of peer interaction within classrooms have been left to further research. Data restrictions have led to the use of increasingly large groups of potential peers, which casts doubt on the causal effect of peers found in the literature. Thus, selecting relevant peers would increase trust in the relationship between peer performance and attributes and the individual outcomes.

For the latter, sociological theory suggests a more dynamic way of peer interactions by focusing on interpersonal relations. For example, considering the likelihood of peer interaction and interpersonal ties in a closed group, such as workplace, class or neighborhood, allows for a description of the social network and the positioning of network nodes. I hypothesize that within a class, dyadic weak tie relations can be postulated based on status and value homophily, i.e., that student's group and cooperate according to both visible and latent characteristics (McPherson et al., 2001). These behaviors are important to determine the importance of friends in the construction of peer averages in order to attain a true impact of peers (Lin, 2010).

This paper is part of a large literature on the production functions of education, particularly cognitive skills. The emphasis is on the modeling of peer effects, which are relevant in the discussion of school choice or voucher programs¹, in the policy discussion of tracking programs, or regarding the permeability of educational domains.

Proceeding from the seminal work of Manski (1993), literature on peer effects aims at the correct identification of peer effect parameters. Simpler models employing average class performance as a measure of peer effects are at risk of not being able to separate pure peer effects from contextual effects, i.e., common exogenous factors such as economic or institutional environment. Accordingly, a large literature emerged dealing with approaches to overcome the *reflection problem* described by Manski. However, most of these approaches rest on the properties of the datasets used, which limits the application to a handful of countries and, more important, forces to make crucial assumptions on the relevant peer relations and the level at which they arise, i.e., by school, grade, neighborhood, or even higher levels of aggregation. Such broad definitions impede the identification of true causal impacts as they incorporate the influence of possibly irrelevant peers (Angrist, 2014).

A handful of studies have chosen to control for peer effects implicitly through the inclusion and interaction

¹E.g., No Child Left Behind (NCLB) act or the Boston Metco program (Angrist and Lang, 2004)

of fixed effects in panel data sets (Arcidiacono et al., 2012; Badland and Schofield, 2006). In a setting of stable classes or groups, the unobservable baseline ability of peers can be captured through the inclusion of student-level fixed effects. However, the majority of large scale student achievement tests, i.e., TIMSS, which is used here, as well as the Progress in International Reading Literacy Study (PIRLS) and the OECD's Programme for International Student Assessment (PISA) as the most prominent examples, is set up as a repeated cross-section, thus rendering fixed effect estimation techniques inapplicable. Instead, these datasets require an explicit modeling of peer relations, for which mostly the average of peer variables is chosen.

Only rather recently, the technique of spatial regression has been employed in the research on peer effects in various domains, e.g., achievement outcomes, recreational activities, or healthy lifestyle (Lin, 2010; Bramoullé et al., 2009; Boucher et al., 2014; Lee, 2006). Spatial models allow for the consideration of social networks, i.e., the degree to which each peer affects the individual is allowed to vary. Thus, the structure of the peer group can be represented as a matrix of social interactions. Again, the aforementioned studies rely on datasets that offer information on friendship structures, e.g., through the closest friends reported in the student survey.

I contribute to the literature by using sociological theory to construct social network matrices that describe classroom structures within the 2015 wave of TIMSS. These social network matrices can be employed in spatial regression approaches to model peer effects in the TIMSS 2015 dataset, even in the absence of explicit information on friendship relations. This strategy allows for smaller groups of relevant peers, which resolves doubt of the positive peer effects not being causal.

I use regression techniques from spatial econometrics, which aim at modeling geographic interdependencies, together with maximum likelihood estimation to model peer effects within a classroom. The explicit modeling of peer relations facilitates the understanding of how classmates affect each other by allowing for a weighting scheme of peer achievement and enables establishing a causal link between the characteristics and traits of relevant peers and individual student achievement. Additionally, the estimation of spatially lagged dependent and independent variables allows for the separation of endogenous and exogenous peer effects, i.e., individual outcomes can be affected by both peer outcomes and peer characteristics. This circumvents the so-called reflection problem in the estimation, described by Manski (1993), which has ever since played a crucial role in the identification of peer relations.

For the 2015 wave of the Trends In Mathematics and Science Study (TIMSS), socio-economic spatial weights matrices are constructed to provide measures of similarity for students in each class following the concepts of *status* and *value homophily*. In an education production function setting, these matrices based on personal network integration and weak-tie relations allow for quantifying the externalities arising in classrooms in secondary education for 8,566 students in 8th grade in the United States. This strategy is supposed to enable for a more causal modeling of peer effects, as the influence is restricted to the relevant peers and it furthermore contributes to the reduction of omitted variable bias in the estimation of determinants of individual student achievement. I include a full set of student-specific variables, comprising individual,

parent, and school-level factors relevant in the production of cognitive skills, in order to separate and correctly identify the different effect levels of peer inputs.

In a preliminary step, I use an ordinary least squares (OLS) or simple linear model (SLM) model to estimate the education production function using only standard inputs, such as individual and parental characteristics, as well as school and environmental factors. I proceed by using spatial regression techniques with a weighting matrix that corresponds to the standard linear-in-means (LiM) model for each classroom. These unweighted peer effects show a significant and positive association with individual level achievement in a spatial autoregressive model. The results indicate that peer (or indirect) effects on key variables have a considerable magnitude and both endogenous and exogenous peer effects can be identified. Furthermore, school level control variables lose significance in the presence of peer effects. This estimation method provides the baseline for the further analyses of peer cooperation behavior.

For this more detailed analysis of which students are more likely to collaborate, I construct two similarity matrices to account for the fact that students with more similarities have an increased likelihood of being in a peer group. The first matrix is based on visible characteristics discussed in the literature (e.g., gender, age, origin) and corresponds to the theory of status homophily. The second weighting matrix further includes student achievement to account for the overall attitude towards learning, as suggested by the concept of value homophily. This also encompasses the basic idea of heterogeneous effects in peer relations, so that influence depends on the positioning in the achievement distribution. Changes in the social network matrices employed leave the qualitative results attained from the prior step unaltered, yet increase the confidence in the causal influence of peer effects by highlighting the role of relevant peers.

In the regression analysis, I employ variables on schooling environment and the school's neighborhood to account for *contextual effects* or environmental factors common to all students. However, as a further sensitivity analysis, the main findings are also robust to a specification in which school-level covariates are implicitly modeled by the consideration of fixed effects on the school level. These capture all remaining, unobserved heterogeneity that is specific to the school, but constant over time, such as socio-economic and cultural environment, quality of teaching staff, or availability of educational resources. However, the main results change only little upon the consideration of the fixed effects.

The remainder of the paper is organized as follows. Section 2 reviews the relevant literature on social interaction and personal network integration of adolescents and highlights its relevance in the estimation of education production functions (EPF). Section 3 outlines the empirical study together with the construction of sociology-based social network matrices and a detailed discussion of the main control variables and their respective transmission channels. Section 4 provides the main results as evidence for the role of peer effects in the estimation of individual level student performance and a comparison to other studies. Section 5 summarizes the main findings and concludes.

2 Theoretical Considerations

Literature on the estimation of cognitive abilities suggests treating the observed educational outcome (here: student achievement test scores) as the result of a production process. Accordingly, an individual's performance can be modeled as the result of employing environmental inputs and idiosyncratic efforts made in the production of knowledge. Equation (1) shows a generic model, with the outcome O_{ict} of individual i in country c at time t as a time-dependent function of cumulative family $F_{ic}(t)$, educational inputs $S_{ic}(t)$, peer effects $P_{ic}(t)$, a time-independent innate ability A_{ic0} and a random error component ε_{ict} that captures all remaining unobservable factors.

$$O_{ict} = O_t[F_{ic}(t), S_{ic}(t), P_{ic}(t), A_{ic0}, \varepsilon_{ict}] \quad (1)$$

For TIMSS, individual level achievement data, as well as consistent information on family background (i.e., socio-economic and educational values) and some school inputs (e.g., location, resources, composition), are available. However, these inputs are measured for the respective year when the study was conducted only and provide no information on cumulative or historical inputs. A proximate measure of innate ability (e.g., results of standardized IQ tests, or prior performance) is also not part of the data. For this case, i.e., the modeling of the level of achievement, [Todd and Wolpin \(2003\)](#) suggest using a contemporaneous specification, although it imposes some assumptions and restrictions on the conditioning variables used.

First, this specification assumes that only contemporaneous inputs are used in the production of current achievement. Second, inputs are assumed to be constant over time, or at least that their contemporaneous level provides a valid approximation of the history of inputs. Third and last, it is required to assume that the contemporaneous inputs are not correlated to any unobserved variables. These restrictions leave the prior equation reduced to a form incorporating only contemporaneous inputs as portrayed in equation (2).

$$O_{ict} = O_t(F_{ict}, S_{ict}, P_{ict}) + \varepsilon_{ict} \quad (2)$$

Any unobserved heterogeneity is captured in the error-term ε_{ict} . If the third assumption holds, this is unproblematic regarding a possible omitted variable bias on the remaining coefficients. However, if there is correlation among the inputs, this will lead to biased and thus inconsistent coefficient estimates. For peer effects, this scenario appears very likely. Consider peer effects being omitted from equation (2) and the aforementioned inputs being relevant factors in determining individual achievement. In this case, they are also determinants of peer performance and thus correlated. If peer effects are disregarded in the estimation of an EPF, all the other regression coefficient estimates are biased as well. Accordingly, it appears advisable to always use peer effects in an EPF to attain unbiased coefficient estimates.

However, upon the consideration of peer effects, a technical difficulty in the estimation of an EPF arises through simultaneity. This problem is best illustrated by a simple example. For an education production function in a class of two students, let O be the educational outcome for students i and j . It is determined by a vector X of relevant characteristics, the educational outcome of the respective other students in the class and an error term. If an exogenous shock occurs for student i (e.g., domestic violence, as is investigated by

Carrell and Hoekstra, 2010), resulting in a fall of O_i , this will also cause a decrease of O_j through i 's effect on j 's educational outcome.

$$O_i = X_i\beta + \lambda O_{-i} + \varepsilon_i \quad (3)$$

$$O_j = X_j\beta + \lambda O_{-j} + \varepsilon_j \quad (4)$$

Here, the fall of the educational outcome of j as an exogenous shock is falsely identified as the cause of the decrease in achievement of i . This leads to the identification of a spurious positive relationship between individual and peer performance, instead of measuring the true effect of peer outcomes. This entails the estimates of the remaining coefficient being biased and results in false inferences and invalid hypothesis testing. As a consequence, considering peer effects would not reduce omitted variable bias, but instead introduce further bias through the simultaneous relation of outcomes.

The prior example highlights another fundamental problem in the discussion of peer effects. For a valid identification of peer influences, the possible effects of peer outcomes and peer characteristics must be separated. The literature on peer effects has prevalently adopted the vocabulary introduced by Manski (1993), who distinguished peer effects depending on the context in which they arise. First, in *endogenous effects*, the behavior of the individual studied is affected by the behavior of the related peers. In this case, the outcome under study is affected by the outcome of the peers, i.e., individual achievement is partly determined by peer achievement. Second, *contextual* or *exogenous effects* measure how exogenous group characteristics affect the way the individual studied behaves. In the education production function, this would correspond to how the inputs of peers (e.g., parental or individual inputs) affect the outcome of the individual studied. Third and last, *correlated effects* are the cause of individuals behaving alike as a consequence of having similar characteristics or being exposed to common environments. These are influences that affect all individuals within the class, i.e., school level characteristics or neighborhood effects.

Transmission Channels Depending on the setting, either peer outcomes as an endogenous effect or peer characteristics as an exogenous effect (or as another possibility, a combination of both) could affect individual outcomes. These considerations of transmission channels are crucial for the correct choice of spatial regression models in the subsequent study. Individual scholastic achievement is in parts determined by the performance or the behavior of other students in the same school or class, or in the group of friends. In economic terms, these peer effects can best be described as externalities in the production of cognitive skills or human capital. In spite of a vast literature on peer effects, it is not directly clear through which of the various possible channels and in which direction peer effects work. Ties between students could serve as transmission channels for material and immaterial goods, or information, especially help with learning or study groups, or simply working together in class. Plentiful proponents of positive externalities provide examples of how high-achieving peers could boost an individuals performance, e.g., by encouraging learning atmosphere and creating incentives to reach a similar level of achievement (Hanushek et al., 2003; Burke and Sass, 2013; Carrell et al., 2009; Arcidiacono et al., 2012). This suggests that efforts to alter school composition (e.g.,

through vouchers or bus transfers) could serve as a valid improvement of overall educational performance. On the contrary however, some peer characteristics could have adverse effects, so that the performance of individuals is dragged down. One possible negative effect is overall low classroom performance, which could slow down the learning progress of the whole class and thus reduce efficacy of teaching. However, this is simply the reverse of the medal of the aforementioned relationship. Other than that, high achieving peers could lead to the teacher having higher expectations, thus grading the bottom-end of the achievement distribution more strictly (Kiss, 2017). Also, high-achieving peers could be intimidating to their low-achieving counterparts and thus inhibit learning efforts (Winston and Zimmerman, 2004). Additionally, some behavioral peer characteristics rather than their performance could affect individual achievement. Other studies find significant and negative impacts of peers who show disruptive behavioral symptoms, such as absenteeism, distraction, or substance abuse. These undesirable traits reduce the effectiveness of the learning environment and furthermore put peers to show similar patterns of disruptive behavior (Grygiel et al., 2018; Carrell and Hoekstra, 2010; Gaviria and Raphael, 2001). These findings illustrate that peer effects can arise in the form of outcomes as well as characteristics. Thus, for the precise modeling of peer effects, both endogenous as well as exogenous peer effects should be considered.

Why are peer effects especially relevant in determining the performance of teenagers in school in the first place? Generally, students in secondary education are regarded as vulnerable to influences of either sort, as they are still forming their character and personality (Steinberg and Monahan, 2007). In this study, students from the TIMSS United States population, with an average age of about 14.2 years, are among the most susceptible adolescents to influence on their behavior from peers in and out of school. Sorensen et al. (2017) analyze specifically how the influence of peers and families evolves over time. Their main finding suggests that from fourth to seventh grade, role models shift and thus the importance of parental influence in educational achievement nearly vanishes, while the effect of peer influences multiplies significantly. This shift is attributed mainly to a change in role models with age, i.e., for teenagers, peers are the ones to determine attitude and effort towards learning and how time is allocated to different activities or hobbies. The magnitude of peer effects is likely to depend on the setting, i.e., class or school size, and on other factors, e.g., socio-economic or cultural background (Carrell and Hoekstra, 2010). The magnitude of the effect also depends on the position in the distribution of peer group achievement. For instance, high-achieving students are found to receive less benefits from average peer performance than low-achieving classmates do (Hanushek et al., 2003). Furthermore, peer effects are likely to influence individual achievement even outside school, i.e., through common leisure activities, working jointly on homework or spending time together in other ways (Knifsend et al., 2018).

The discussion of peer groups of course raises the question of how students are in fact grouped within the class and how they choose to cooperate. On the one hand, groups could be ex ante determined by the educator in order to optimize the learning environment, while on the other hand, self-selection and internal building of closed groups or cliques is also feasible - as well as a combination of both (Farmer et al., 2018; Freeman et al., 2017). When able to group freely, students are observed to arrange by prior

achievement, ability or interest in the topic. Thus the stereotype of the diligent 'nerd' sitting in the front row and the underachieving 'lazybones' slouching in the back row of the classroom applies to a certain degree. Furthermore, gender and socio-economic or ethnic characteristics are found to contribute to the early decision of group formation. Undergraduate students have furthermore been observed to group and collaborate based on especially obvious characteristics such as their ethnicity and gender (i.e., to some extent driven by a stereotype threat to avoid isolation in class), but also their academic history regarding grade average or the attending of common classes (Freeman et al., 2017).

This tendency to join or work in rather homogeneous groups indicates that the sociological theory of *homophily* applies to classroom arrangements and extends to further environments out of school (Louch, 2000; Knifsend et al., 2018). As a general concept for interpersonal ties, a distinction between strong and weak tie relations can be made to describe the levels of interaction in situations, such as workplace, school, family or leisure (Granovetter, 1973).

As a consequence of proximity in a classroom, it is likely that all class members interact in some regard. For the classroom atmosphere and externalities occurring, it is of greater relevance to analyze group dynamics and weak links between classmates than strong tie relations among cliques or closed groups of friends. However, to establish causality of peer effects, it is necessary to identify the relevant peers that actually influence the behavior of the individual. The majority of studies fails to do so (Angrist, 2014), by either attributing the same influence to all peers (e.g., in the Linear-in-means model, Manski, 1993) or by making *ad hoc* assumptions on peer relations (e.g., the Bramoullé et al. (2009) *parade* of peers lined-up, thus making possible a lag-structure to identify *nearest neighbors* among classmates). Only few studies provide actual information on friendship statuses reported by individuals under study (e.g. Lin, 2010), which however come with restrictive datasets and missing information on peer relations.

Endogeneity of Schooling Location and Class Composition The consideration of peer group effects becomes important to parents and policymakers when school choice, access, integration, or tracking decisions are made (Epple and Romano, 1998; Angrist and Lang, 2004). While proponents of tracking systems argue that homogeneity in schools and classes increases the efficiency of education, opponents favor potential benefits of heterogeneous groups, where high-performers boost motivation and quality of the learning environment (Hanushek et al., 2003; Hanushek and Woessmann, 2006). Especially parents are anxious about their children being grouped with low-achieving, disruptive students in the case of achievement-related tracking (Carrell and Hoekstra, 2010; Grygiel et al., 2018). The level of peers regarding achievement and background then cannot be treated exogenously, but is, at a certain cost and timing of adjustment, determined endogenously. The final outcome of peer group formation is *a priori* influenced by the composition and stratification of the population to be clustered, i.e., the pool of potential peers. This is referred to as *propinquity*, which implies that geographical proximity and frequency of interaction increase the likelihood of forming ties (Moreland and Beach, 1992). Conclusively, dyadic homophilous relations that emerge within the class cannot be attributed to *choice homophily* alone, but are mostly due to the group composition. This corresponds to what

McPherson and Smith-Lovin (1987) describe as *induced homophily*. It is thus not possible to completely separate the effect of class heterogeneity and individual homophily, so that homophilic selection occurs in initially relatively homogeneous groups. Accordingly, perceived peer effects might simply reflect a beneficial common environment (e.g., socio-economic, infrastructural, or institutional), entailing false conclusions and biased estimates.

In an attempt to avoid correlated socioeconomic backgrounds of students, Sacerdote (2001) exploits the property of random assignment of college room- and dormmates to circumvent the presence of correlated effects. If the assignment is truly random, decisions on social interaction can be understood as a result of *choice homophily*. However, the sample being drawn from the population of a US college indicates that similar backgrounds are more than likely and the assignment of peers is not pure random, but instead based on the prior selection through the college admission process, thus rather reflecting *induced homophily*. Another approach to control for this issue technically is made by Ammermueller and Pischke (2009) who employ a model of school or even teacher fixed effects in a study of peers in elementary schools to capture the larger part of contextual effects. In their model, peer effects arise in the form of exogenous effects, which, as they acknowledge, might also be driven by endogenous effects. Accordingly, their approach is valid for the identification of peer effects in general, yet not suitable for the distinct identification of exogenous and endogenous effects. Manski (1993) describes the non-random formation of social networks as uncritical, if it is applied in studies of small-group interactions. The assignment of students to a class implies that all individuals know each other to a certain degree (e.g., information on performance, attitude, or background is available). Within each class, students make decisions on social network formation based on the remaining class heterogeneity and choose their efforts accordingly after having been selected to the sample.

3 Empirical Approach

In order to validate the suitability of peer relations postulated on the base of weak tie relations, I run a spatial regression approach using individual level student achievement as a dependent variable. The purpose of the study is to confirm the presence of peer effects in secondary education and to show the suitability of sociological theory in the modeling of peer relations and classroom composition.

3.1 Estimation Sample

The study uses the data on student achievement in 8th grade of the USA from the International Association for the Evaluation of Educational Achievement (IEA) current wave of TIMSS from 2015. The TIMSS dataset provides information on the school and class membership of each individual, so that it is possible to group students by their class, instead of drawing peer relations from grade-level or school populations. Each class constitutes the population from which individuals draw the members of their peer group. This data structure ensures that peer effects measure actual relations which fulfill the criterion of real exposure or *propinquity*, i.e., geographical proximity (Moreland and Beach, 1992), thus indicating a causal impact of possible peers (as opposed to the concerns addressed in Angrist, 2014).

In total, the dataset provides information on 8,566 individuals. An important feature of the TIMSS dataset is that information is available, how the students are distributed among the 213 schools and 457 classes in the sample. Compared to prior studies, this allows for credible class sizes and thus reasonable peer groups. The minimum, mean, median and maximum class size are 1, 18.7, 20, and 37, respectively, whereas the distribution is shown in figure 1. The sample comprises 6 isolated individuals, for whom no peers could be assigned as they have a unique school and/or class id. In total, 59 individuals are reported as being in classes with 5 or less students, representing 0.69 per cent of the sample population.²

TIMSS encompasses schools of all types (public or private), as well as all common educational tracks in the US education system. The students in the estimation sample have an average age of 14.23 years and consist of 50.52% girls and 49.48% boys. 94.50% of the students have stated that they were born in the territory of the United States.³

3.2 Socio-Economic Spatial Weights Matrices

Adjacency Matrix As a precondition, to ensure that individuals are only affected by students of the same class, a binary contiguity social network spatial weights matrix, the adjacency matrix is generated. If two students i and j are enrolled in the same class, they are interpreted as peers, so that the element of the matrix w_{ij} equals 1. Of course, when one class is considered, all class members are peers to each other, except for themselves, so that $w_{ii} = 0$. Accordingly, the classroom spatial weights matrix is a symmetric square matrix with 0 as diagonal and 1 as off-diagonal elements, as shown in equation (5).

$$\mathbf{C}_c^{(Adj.)} = (w_{ij})_{n_c \times n_c} = \begin{pmatrix} w_{11} & w_{12} & \cdots & w_{1n_c} \\ w_{21} & w_{22} & \cdots & w_{2n_c} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n_c1} & w_{n_c2} & \cdots & w_{n_cn_c} \end{pmatrix} = \begin{pmatrix} 0 & 1 & \cdots & 1 \\ 1 & 0 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 0 \end{pmatrix} \quad (5)$$

This spatial weights matrix is row normalized, i.e., each row is divided by the number of non-zero elements. The direct interpretation of these matrices is that in a first step, all classes are analyzed in isolation and each student of the same class is considered as a possible peer, whereas students from other classes, even in the same school, do not qualify. To create the final social network matrix for the estimation sample, the class matrices C_l^c with $c \in [1, \dots, C]$ being the number of each class, are aligned on the diagonal of a main matrix, where all other elements are zero, as is depicted in equation (6).

²As opposed, the sample used in the study by Lin (2010), contains 68,131 observations, of which 18,572 are isolated. Those are individuals which have not been reported as a friend by any other person in the data and have also not reported any friends from the dataset. For the remaining observations, the minimum, mean and maximum group size are 2, 38.5, and 427 observations, respectively.

³According to supplement 2 of the TIMSS dataset, "United States" includes the 50 states, its territories, the District of Columbia, and U.S. military bases abroad"

$$\mathbf{W}^{(Adj.)} = \begin{pmatrix} \mathbf{CI}_1^{(Adj.)} & \cdots & \mathbf{0}_{n_1, n_c} & \cdots & \mathbf{0}_{n_1, n_C} \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ \mathbf{0}_{n_c, n_1} & \cdots & \mathbf{CI}_c^{(Adj.)} & \cdots & \mathbf{0}_{n_c, n_C} \\ \vdots & \cdots & \cdots & \ddots & \vdots \\ \mathbf{0}_{n_C, n_1} & \cdots & \mathbf{0}_{n_C, n_c} & \cdots & \mathbf{CI}_C^{(Adj.)} \end{pmatrix} \quad (6)$$

This specification of the social network mathematically equals an averaging of the peer attributes across all potential peers, which is consistent with a traditional modeling of within-class peer effects. In terms of the [Manski \(1993\)](#) vocabulary, this corresponds to a linear-in-means model, where achievement is determined by the linear individual level inputs and the mean achievement and inputs of the peer group. Accordingly, every student in a class is affected by the mean performance of all classmates.

Similarity Matrices To quantify the likelihood of dyads being peers instead of just analyzing a closed group in total, it is necessary to evaluate the proximity of all pairs of individuals within a class. [Granovetter \(1973\)](#) names interaction, affection and time as the preconditions for strong tie relations. From the TIMSS data, the first two are not observable while the last is only partly measured by common class membership. This proximity could be a driver of the bonding process of peers ([Moreland and Beach, 1992](#)). Conclusively, the data allows for the analysis of weak tie relations only, based on individual level characteristics (gender, age, origin) and attributes (i.e., performance) provided.

Similarity is measured based on a coefficient by [Gower \(1971\)](#), which is designed for mixed measurement scales of variables⁴. Consistent with the theory of weak tie relations and the concept of *homophily*, Gower similarity indicates the likelihood of a dyadic relationship. Accordingly, all dyadic relationships portrayed in the matrix are symmetric, i.e., tie $(i \rightarrow j) = (j \rightarrow i)$.

$$S_{ij} = \sum_{k=1}^{\nu} s_{ijk} \delta_{ijk} / \sum_{k=1}^{\nu} \delta_{ijk} \quad (7)$$

To incorporate mixed measurement scales, the computation of s_{ijk} differs according to the variable scale at hand. For quantitative variables, the similarity measure is computed as given in equation (8), where R_k is the range of the variable considered. The parameter δ_{ijk} represents the value 1 when a character k can be compared and is set to 0 otherwise. The formula implies that for each character, the maximum similarity is 1 for the case that $x_i = x_j$ and the minimum similarity is always zero between the lowest and the highest value of the specific character.⁵

$$s_{ijk} = 1 - |x_i - x_j| / R_k \quad (8)$$

⁴E.g., while *age* is a metric variable, both *gender* and *migration background* are binary variables

⁵This would correspond to $s_{max,min,k} = 1 - |x_{max} - x_{min}| / R_k$, where $R_k = x_{max} - x_{min}$, so that $s_{max,min,k} = 1 - \frac{|x_{max} - x_{min}|}{(x_{max} - x_{min})} = 0$

In the case of dichotomous variables, s_{ijk} and δ_{ijk} are defined according to the rule presented in (9). The similar approach is used for qualitative data, where $s_{ijk} = 1$ if individuals i and j agree on the k -th character and is set to zero otherwise.

$$s_{ijk} = \begin{cases} 1 \wedge \delta_{ijk} = 1 & \text{if } x_i = x_j = 1 \\ 0 \wedge \delta_{ijk} = 1 & \text{if } x_i \neq x_j \\ 0 \wedge \delta_{ijk} = 0 & \text{otherwise} \end{cases} \quad (9)$$

By using relationship likelihood rather than observed relations, I introduce some assumptions on social network characteristics. First, for network density, no maximum number of connections or relations exists. All students affect each other, although the less similar have a smaller effect and vice versa. Second, no closed clusters exist within classes. This implies that no formation of closed cliques or groups occurs and that peer membership does not affect the likelihood of relationship to a non-member of the specific peer group. Third, transitivity in dyadic relations⁶ is disregarded. Fourth, the multiplexity of the relationship cannot be assessed. Thus, it is not possible to determine at what level a relationship occurs (e.g., friendship, comradeship, or strategic cooperation). Fifth, and last, the constructed similarity matrix above is symmetric and strictly positive, i.e., it corresponds to the description of an undirected network, where the direction of influence between two nodes is disregarded. Furthermore, all peer relationships are beneficial, so that the possibility of enmity between students is excluded by definition

Compared to the baseline adjacency matrix described above, the similarity measures based on the Gower coefficient introduce a weighting of peer influences. Depending on the specification, individuals experience stronger influence from peers with similar characteristics. This does not imply that, e.g., boys cannot influence girls as peers, but rather that this is less probable.

Status Homophily Matrix To express the tendency of students to form peer relations according to visible characteristics, I construct the first similarity matrix based on the student's gender, age and nationality (Carrell and Hoekstra, 2010; Diette and Uwaifo Oyelere, 2017). As these factors can be considered as fixed, they are subsumed under the term of *status homophily*. This setting implies that peer groups emerge in the early phase of class assignment, based on rather superficial characteristics, with the possible aim of reducing the initial risk of becoming an outcast (Freeman et al., 2017; McPherson et al., 2001).

⁶E.g., if $a \leftrightarrow b$ and $b \leftrightarrow c$, then $a \leftrightarrow c$

$$\mathbf{CI}_c^{(SH)} = \begin{pmatrix} S_{11} & S_{12} & \cdots & S_{1n_c} \\ S_{21} & S_{22} & \cdots & S_{2n_c} \\ \vdots & \vdots & \ddots & \vdots \\ S_{n_c1} & S_{n_c2} & \cdots & S_{n_cn_c} \end{pmatrix} - I_{n_c} \quad (10)$$

This classroom similarity matrix is denoted as $\mathbf{CI}^{(SH)}$ and consists of the dyadic Gower similarity measures for all students, based on the variables gender, age, and whether the student was born in the United States. The matrix is square and symmetric, i.e., all relations between the class members as the nodes of a social network are undirected. The Gower similarity measure S_{ij} in the case $i = j$ would correspond to 1, as no difference on any factors can be observed. This violates the precondition that nodes in a social network matrix do not affect themselves, thus in equation 10 all elements on the diagonal set to zero by subtracting an I_n matrix. Subsequently, the matrix W^{SH} is again row normalized.

Value Homophily Matrix While the prior matrix is based on visual characteristics and corresponds to a rather short-term formation of groups, it is possible to exploit non-visual characteristics for the evaluation of similarity. Hanushek et al. (2003) list achievement, in addition to race and socio-economic status as the three most commonly used characteristics in the separation of peer group effects. Including achievement as a measure on which peer relations are chosen serves two purposes. First, students have been observed to group according to academic performance (Freeman et al., 2017) and second, prior results suggest that peer effects are heterogeneous for different positions in the achievement distribution (Hanushek et al., 2003; Sorensen et al., 2017). Accordingly, the value homophily social network matrix can be interpreted as a rather long-run formation of peer groups, where students have demonstrated or proven their academic performance. The computation is identical to the status homophily matrix and relies on Gower’s coefficient of similarity as well.

3.3 Estimation

In order to evaluate individual student achievement, I estimate an education production function as specified in equation (11). In order to consider the effect of peers and classroom structure, peer effects are explicitly modeled in the regression, using spatial regression techniques based on different socio-economic spatial weights matrices to represent sociological concepts of the formation of social groups.

$$O = \alpha \nu_n + \rho \mathbf{W}O + \mathbf{X}\beta + \mathbf{W}\mathbf{X}\lambda + \mathbf{P}\mathbf{a}\delta + \mathbf{S}\mathbf{c}\gamma + \varepsilon \quad (11)$$

where $\varepsilon \sim N(0, \sigma^2 I_n)$

Here, O represents the educational outcome of student i in school s , and class c . \mathbf{X} refers to a matrix of individual characteristics, while the matrix $\mathbf{P}\mathbf{a}$ measures parent level inputs in the education production process that are, of course, specific to the individual. To account for the school and classroom environment,

i.e., the *correlated* effects, a matrix of school characteristics (\mathbf{Sc}) is included.

Additionally, \mathbf{WO} refers to the influence of classmates on the individual performance, where \mathbf{W} is a socio-economic spatial weights matrix, which can have various definitions (see above), and O is a matrix of student achievement. The coefficient ρ is the *endogenous* peer effect. Similarly, \mathbf{WX} refers to spillover effects of peer characteristics on the individual, where \mathbf{X} denominates a matrix of student characteristics. This corresponds to the *exogenous effects* described above. Finally, the variable term ε is a stochastic error term.

As [Manski \(1993\)](#) and [Lyle \(2007\)](#) add for consideration, OLS estimation of peer effects is likely to be biased due to the *reflection* problem induced by simultaneity⁷. Technically, using the contemporaneous average achievement of peers as an explanatory variable corresponds to an auto-regressive component across individuals. OLS estimation of this relationship would introduce a spatial dependency into the error term. This leads to a biased coefficient of peer achievement and henceforth to biased estimates of all other coefficients as well. An alternative estimator that is capable of considering the network interdependencies in the dependent variable of students within the class is the maximum likelihood (ML) estimation. The estimation is based on the assumption of normality in the error term, which, for the i.i.d. case, simplifies to $\varepsilon \sim N(0, \sigma I_n)$. This disturbance term cannot be directly observed and is instead based on the dependent variable O .

$$\varepsilon = (I_n - \rho W)O - (\alpha I_n + \mathbf{X}\beta + \mathbf{WX}\lambda + \mathbf{Pa}\delta + \mathbf{Sc}\gamma) \quad (12)$$

In order to solve the vector ε into O , the Jacobian term corresponds to equation (13).

$$|\partial\varepsilon/\partial O| = |\partial[(I_n - \rho W)O - Z\delta]/\partial O| = |(I_n - \rho W)| \quad (13)$$

The likelihood function for the ML estimation of ε is

$$L(\varepsilon|\theta) = \frac{1}{2(2\pi\sigma^2)^{n/2}} \exp\left(\frac{-1}{2\sigma^2}\varepsilon'\varepsilon\right) \quad (14)$$

By multiplying the likelihood function for the disturbance term ε in equation (14) with the Jacobian in eq. (13), the likelihood function for the observable O is obtained

$$L(O|\theta) = L(\varepsilon|\theta) \times |\partial\varepsilon/\partial O| \quad (15)$$

For the log likelihood function for the spatial lag model, equation (15) is logarithmized, as is portrayed in eq. (16)

$$\mathcal{L}(O|\theta) = -\frac{n}{2}\ln(2\pi\sigma^2) - \left\{\frac{n}{1}(2\sigma^2)\varepsilon'\varepsilon\right\} + \ln|(I_n - \rho W)| \quad (16)$$

⁷i.e., the individual performance affects mean class performance, yet mean class performance affects the individual performance

This equation only depends on the parameter vector θ and can thus be solved numerically to find its maximum and thus yields the maximum likelihood estimator.

3.4 Variable Selection

Various socio-demographic variables that might explain differences in student achievement are included in the regression to account for individual heterogeneity.

3.4.1 Student Characteristics

Gender First, the reported sex of each student is included to control for gender-specific differences in classroom performance and attitudes towards schooling. Reportedly, male and female students differ in terms of diligence, ability and achievement, depending on the subject considered, presumably due to differences in rearing and cultural aspects (Guiso et al., 2008; Fryer and Levitt, 2010; Aesaert and van Braak, 2015) or due to gender-specific curricular needs, e.g., physical activity as a balancing factor for boys in order to increase the ability to concentrate and stay focused (Cöster et al., 2018).

Age As TIMSS is designed as a grade specific study and the focus is on 8th grade, the age of the observed students is not fixed⁸. Even though a minimum age for enrollment is mandatory in many US states, there is some scope in the parental decision for enrollment (with required entry between 5 to 7 years), which causes significant age-differences of first-graders with diverse affects on educational outcomes (Bedard and Dhuey, 2006).

Students could be young relative to their grade as a consequence of *acceleration*, i.e., by early enrollment or skipping of a whole grade. This way, young age for grade could be caused by high initial abilities, but also indicate a lack of maturity, mentality, or experience, entailing impaired classroom performance or social outcomes in later life and is thus discussed controversially (McEwan and Shapiro, 2007; McCrary and Royer, 2011). Similarly, the effect of skipping a grade is considered to be ambiguous. High achieving students who skip grades are observed to benefit from their new peers in an academic sense, as well as in subsequent earning possibilities (Warne and Liu, 2017; Warne, 2017). However, they suffer from socio-emotional deficiency and difficulties in finding compatible peers (Kretschmann et al., 2016).

As opposed, high age relative to grade, could be the consequence of either voluntary late enrollment or of grade repetition. Typical reasons for the latter are bad grades, lack of motivation or participation, or absence due to serious disease or important reasons (Hughes et al., 2017). For developed countries and emerging markets⁹, often the opposite is observed, i.e., that relatively older children show more readiness for school and thus outperform their younger classmates. Thus, in a setting of adequate health and infrastructure,

⁸Different from, e.g., OECD's PISA study, which is targeted to a 15 year old student population which again differs in terms of grade

⁹Specifically, Bedard and Dhuey (2006) analyze the age-premium for a sample based on the TIMSS-1995 participating countries

the gains in maturity could dominate and lead to a positive effect of older relative age ([Bedard and Dhuey, 2006](#); [McEwan and Shapiro, 2007](#)).

Immigrant Status Additionally, a student’s immigration background is included to account for possible difficulties in the learning environment. This is measured by the student’s migration status (i.e., first or second generation immigrant) and the frequency at which the English language is spoken at home.

Several studies have found a migrant background to significantly interfere with the academic career. The effect is not exclusively caused by differences in socio-economic background, but also reinforced by the characteristics of the educational system in the country of residence ([Borgna and Contini, 2014](#)). For instance, in a selective, public tracking system, immigrant children and children from ethnic minorities have been observed to be hampered from transition to secondary education ([Jackson et al., 2012](#)) and are more frequently assigned to non-prestigious educational tracks ([Cebolla-Boada, 2011](#)). However, immigrants are found to take more ambitious decisions in their educational career than their native counterparts, specifically in tertiary education ([Kristen et al., 2008](#)). Furthermore, it has been observed that the performance of natives is negatively affected by the presence of immigrant pupils in the class. Thus, being foreign might affect not only the individual, but also creates negative classroom externalities ([Brunello and Rocco, 2013](#); [Diette and Uwaifo Oyelere, 2017](#)).

Perceived bullying To account for the effect of classmates harassing or bullying on the achievement of the respective student, an item for perceived bullying activity is included. Bullying has been identified as a determinant of student achievement that has significant negative effects on the outcome, regardless of the grade level or the country ([Ponzo, 2013](#); [Resende Oliveira et al., 2018](#)). However, causality could run in both directions, so that initially high performing or ambitious students could even be confronted with increased bullying ([Bekiari et al., 2017](#); [Harel-Fisch et al., 2011](#)).

3.4.2 Family Background

Educational Achievement In order to measure the overall socio-economic status, as well as the cognitive abilities of the individual’s mother and father, the highest educational attainment for each parent is included to provide a measure of initial or innate ability. The socioeconomic status and educational level of parents are generally agreed to have a positive correlation with the school performance of their children ([Martins and Veiga, 2010](#)), with child health as a possible transmission channel ([Currie, 2009](#)). Also, a higher level of education is assumed to increase the valuation, involvement, and the expectations for schooling and education ([Rindermann, 2007](#); [Fan and Chen, 2001](#); [Davis-Keane, 2005](#)). Additionally, a strong intergenerational transmission of intelligence is observed, parent cognitive ability is found to be a strong predictor of the child’s educational outcomes ([Anger and Heineck, 2010](#)).

Immigration Background Similar to the migrant status of the child, a variable indicating whether each parent was born in United States is included. This allows to identify cases in which a native-born citizen

of a country has (partly) foreign-born parents or other combinations. Accordingly, an approximation of the family’s migration history can be made to determine whether the child is a first or second generation immigrant.

Proxies for Home Economic and Educational Resources The TIMSS dataset does not explicitly contain a measurement of the parents economic resources available (e.g., income, wealth, real estate) which could affect the childrens’ educational outcomes (Martins and Veiga, 2010). Instead, I proxy for resources by the presence of items specific to education or wealth at home. For instance, the number of books at home is described as a measurement of parenting, home education, priority for knowledge, and social background (Hanushek and Woessmann, 2011) and is found to be robustly associated with student achievement (Schuetz et al., 2008; Freeman et al., 2010). With a similar interpretation, I include the number of digital devices at home. In the supplement to the TIMSS dataset, these are listed as the possession of a computer or tablet, internet connection, own mobile phone, or a gaming system. These items are thought to proxy for wealth, but also for adjustment to modern technology, i.e., whether the parents have already adapted themselves as well as their parenting style to the requirements of technology-rich environments (Hanushek et al., 2015) and enable their children to do so as well. As a measure of family resources devoted to their children, dummy variables for having an own desk, an own room, or both, are included. This measures the basic prerequisites for the child to recapitulate and review school material at home. Especially for low-income adolescents, the role of parenting style is emphasized (Johnsen et al., 2018).

3.4.3 School Inputs

Location To control for the property of different schools within the United States their location is used. First, the location is measured by a description of the area where the school is situated, i.e., whether it is densely populated, suburban, a medium size city or large town, a small town or village, or remote rural. Second, the amount of people living in the area is included as a second measure. It is expected that not only the characteristics of the school itself, but also further educational and recreational opportunities are affected by where the education facility is located, i.e., whether it is situated in an urban, densely populated area, or rather in a small or remote rural village.

A school’s location is an important determinant of its socio-economic environment and thereby of the composition of students. It has been observed that high-skilled workers cluster in metropolitan areas, whereas rural areas are predominantly inhabited either by old people or those engaged in low- to medium-skilled occupations (Berry and Glaeser, 2005; Moretti, 2013). With parental inputs being gradually higher in larger cities as a consequence of higher average education there, the aggregate school level achievement should benefit from these parental inputs.

Furthermore, large cities could profit from economies of scope. Higher population density enables for a more efficient provision of public goods (e.g., schools, public libraries, or digital services (De Witte and Geys, 2011; Bekkerman and Gilpin, 2013)) which are beneficial for the creation of cognitive skills, with effect size decreasing with distance (Bhatt, 2010).

Last, the location of a school might also be related to the access to recreational or leisure activities and outdoor pursuits. Proximity to schools generally allows for participating in extra-curricular activities, which are found to be beneficial for school performance by promoting a sense of belonging (Knifsend et al., 2018). These activities provide a balance to learning environments and promote health-conscious behavior (Wechsler et al., 2000). Nevertheless, the overall activity level of students is found to be lower in larger cities, indicating a non-linear effect (Badland and Schofield, 2006).

Economic environment Additionally, proxy variables for the economic environment are included to control for possible neighborhood effects. Specifically, from the school questionnaire of the TIMSS dataset, the principals have provided information on what percentage of students approximately comes from either *economically affluent* or from *economically disadvantages homes*. School poverty is accepted as a negative influence on educational achievement and is found to be negatively correlated with parental education. Accordingly, schools with higher levels of poverty are found to have a low socio-economic status peer population, with an effect possibly mediated by the personality trait of agreeableness (Nieuwenhuis, 2018). Generally, a higher share of immigrants in the student population has been found to have an adverse impact on student performance (Brunello and Rocco, 2013). It has been observed that foreign language learners differ concerning their learning style and their expectations regarding teacher behavior and thus pose a challenge for teachers (Tasdemir and Yalcin Arslan, 2018). Diette and Uwaifo Oyelere (2017) investigate the peer effect of English language learners on their classmates and try to separate the negative effects of lacking language skills and of ethnicity on student achievement. Overall, a higher percentage of foreign language students within a class or school is thought to reduce student performance by disrupting the learning atmosphere.

The quality of educational facilities can partly be derived from reported shortage of course material or educational resources in general, as is reported by the school principal for both subjects, mathematics and science. The shortage of supply is thought to create unfavorable learning conditions and thereby reduce teaching efficiency.

School attitude The severity of disruptive behavior impairing the learning outcomes of students is measured by the reported frequency of discipline problems. These incidences affect other students either through distraction or overall reduced classroom performance (Carrell and Hoekstra, 2010; Grygiel et al., 2018). I include this variable to rule out the possibility that peer effects capture disruptive behavior as a common factor among all students of a class.

Last, to consider overall school attitude towards performance of the students, the perceived emphasis on academic success, as reported by the principal, is included.

4 Main Empirical Results

OLS Results Table 3 presents the first results on the determinants of individual student achievement, where column (1) shows the OLS estimates and columns (2-10) show the possible spatial models incorporating peer effects implemented with varying socio-economic spatial weight matrices. A decomposition of the effects of the individual level variables into direct, indirect, and total impacts is presented in table 4. All model specifications include a constant as well as parent and school factors. The first column shows the basic OLS model with no modeling of peer effects. The estimates of the regression coefficients are according to the expectations and consistent with prior literature. A gender achievement gap, which is significant at the 1% level, is found for boys, who thus on average outperform girls in mathematics and science by 10.7848 points. The effect of student age is found to have a negative impact on student achievement. Being one year older is associated with a decrease in test scores by 15.1701 points, again significant at the 1% level. All three categories of the frequency of language use are found to be statistically insignificant. Thus, *not always* using English language at home does not reduce test performance in mathematics and science. However, the effect of not being born on the territory of the United States is statistically significant at the 5% level and negative, amounting to a 7.7286 point decrease of test performance. Last, the effect of perceived bullying is estimated to be positive and highly significant, which at first glance appears puzzling. However, this could capture the effect of "being a nerd", i.e., being bullied exactly because of high school performance.

Linear-in-Means Model In columns (2-4), the estimates for the three spatial regression models based on the adjacency matrix, which corresponds to an unweighted averaging of peer effects, i.e., the linear-in-means model, are reported. As a consequence of the social network connections, coefficients in spatial models cannot be interpreted as marginal effects, but have to be instead decomposed into direct, indirect, and total effects. These partitioned effects are derived from the connections between the individuals, portrayed by the social network matrix, and correspond to taking derivatives for the outcome of student i with respect to the inputs of all other students $j \in [1, \dots, n], i \neq j$. A detailed description on the calculation of partitioned effects is provided in the technical appendix, in section B.1.

Column (2) shows the spatial autoregressive (SAR) model, which includes only the endogenous peer effect (i.e., peer achievement outcome) and restricts the influence of peer characteristics to be absent. The spatial lag parameter, i.e., the endogenous peer effect, is estimated to be positive and significant at the 1% level. It shows a rather high spatial correlation of 0.6874 and indicates that peer achievement provides a positive spillover effect on individual performance. The prior results are affected as follows. The gender gap between boys and girls increases to 11.9304 points and is estimated more precisely. This effect is even increased by local spillovers to 12.901 if the direct effect is considered. However, the indirect effect, i.e., the global spillover is much larger by comparison. If all other classmates were boys, the achievement of the individual would increase by as much as 38.152 points, significant at the 1% level. The estimated negative coefficient for student age decreases in magnitude to -10.9742 by almost one third. Again, the direct effect is more pronounced (-11.867). Likewise, the indirect effect of age is roughly twice the size. If every classmate was

1 year older, individual achievement would be reduced by 23.227 points. Further, the indicator variable for being born outside the United States loses significance. The effect of perceived bullying decreases considerably to 1.0247, yet remains positive and statistically significant at the 1% level. Likewise, the direct effect is slightly larger, while the indirect effect is again twice as big. The SAR model based on the adjacency matrix already provides a much better overall fit than the OLS model, as is indicated both by the pseudo R^2 which increases from 0.327 to 0.402 and the log-likelihood, which likewise rises from $-47,556$ to $-45,569$ ($p < 0.01$, $df = 1$).

In column (3), the results for the spatially lagged in X (SLX) model based on the adjacency matrix are presented. This model rests on the assumption of the absence of endogenous peer effects, i.e., individual outcomes are only affected by the characteristics of peers, while the parameter ρ is restricted to zero. Accordingly, the spillover effects only arise via the peer characteristics, yet no feedback mechanisms are at work. The coefficient estimates of the SLX model are fairly close to those of the OLS model in column (1). Only the individual effect of being born outside the United States decreases both in magnitude and significance, and amounts to a 5.4397 points reduction in overall achievement. For the peer characteristic (exogenous peer effects), a statistically significant and positive impact is estimated for those who *almost always* speak English as the language at home. This finding is suggestive of diversity in terms of some bilingualism being beneficial towards educational performance of the class as a whole, thus posing a positive externality. Nevertheless, the exogenous peer effect of being born outside the United States is negative and significant at the 1% level. If all peers were born abroad, the performance of the individual would, on average, decrease by 80.086 test score points. Overall, the SLX model does not provide an improved fit compared to the OLS model, as can be seen from a very modest increase in the pseudo R^2 to 0.333, although the log-likelihood ratio test indicates a significant improvement ($p < 0.01$, $df = 4$).

The spatial durbin model (SDM) presented in column (4) is a combination of the prior two cases and allows for the presence of both endogenous and exogenous peer effects. The endogenous peer effect, i.e., the spatial lag, remains significant at the 1% level and unchanged in magnitude. Also, the estimated regression coefficients of the model are very close to the SAR model in column (2). For all variables, the results are virtually identical with only minor changes in size, yet not in the overall sign or significance. Accordingly, the direct effects are also close to the SAR case with relevant changes occurring only for the indirect effect, which now incorporate peer characteristics. The point estimate for peers being born outside the United States decreases considerably to less than one third, but remains statistically significant at the 1% level. Under the assumption of all peers being born outside the US., a negative spillover effect of -75.926 points would be exerted on the individual. The exogenous effect of peers *almost always* speaking English at home turns insignificant with a heavily reduced regression coefficient. Compared to the SAR model, the SDM model provides only a negligible increase in pseudo R^2 from 0.402 to 0.406. Also the increase in log-likelihood is modest $ll = -45,565$, yet it still constitutes a significant increase ($p = 0.04$, $df = 4$).

Status Homophily Columns (5-7) show the estimates for the three spatial regression models based on the status homophily social network matrix, which introduces a weighting scheme to peer influences based on similarity on the attributes gender, age, and migratory background. This corresponds to a short-term organization of groups within the class, especially in the early phase of newly composed classes.

Column (5) presents the estimation results for the SAR model. The spatial autocorrelation parameter, which is estimated to be 0.6865, is virtually identical to the prior case, also regarding significance. In comparison to the coefficient estimates based on the adjacency matrix in column (2), the estimated gender gap is considerably smaller with the gap between boys and girls being 10.6090 points. This magnitude is comparable to the OLS results and is also significant at the 1% level, yet it is estimated more precisely, as the reduced standard error indicates. The direct effects are again slightly larger due to the feedback, so that the gender gap widens to 11.480 points. The indirect effect is again approximately twice as large. The coefficients for the remaining individual level variables are almost identical and show no major changes as compared to the LiM model. Conclusively, the remaining direct and indirect effects show a similar pattern regarding size and sign. The goodness of fit of the model is not improved as compared to the LiM specification, as indicated by the pseudo R^2 which is slightly reduced (0.401) and the log-likelihood of $ll = -45,575$, which is also smaller than in column (2).

In column (6) the SLX model for the status homophily matrix is presented. Again, the changes in the individual level variables are negligible, as compared both to the LiM's SLX (3) and the OLS model (1). For the exogenous peer effects, minor changes in the estimated regression coefficients for the peer characteristics can be recognized. The externality of speaking English language at home *almost always* is reduced in size, but remains significant at the 1% level. On the other hand, the adverse externality of being born outside the United States increases in size. Based on the pseudo R^2 and the log-likelihood, no clear distinction between the models presented in column (3) and (6) can be made.

Again, the SDM presented in column (7) represents a combination of the prior two models. The endogenous peer effect is negligibly larger and more precisely estimated than in the LiM model. In contrast to the model in column (4), the gender achievement gap is smaller, with boys outperforming girls by just 10.6230 points, which is even smaller than in the OLS model. Feedback effects increase the direct effect of being a boy to 11.484 points while the indirect effect amounts to 22.154 points, both significant at the 1% level. Also, the negative effect of student age is slightly less pronounced, as can be seen from the regression coefficient estimate and the direct and indirect effects. For the frequency of language use, no change in the significance of the estimated regression coefficients and partitioned effects can be observed. The coefficient for perceived bullying remains positive, significant, and unchanged in magnitude. However, the overall fit is slightly worse, as indicated by a Pseudo R^2 of 0.405 and a log-likelihood of $ll = -45,570$.

Overall, the models based on the status homophily social network matrix provide no increase in the goodness of fit as compared to the standard LiM model, but still provides a considerable improvement to the standard OLS specification. Additionally, the basic idea of assigning more weight to peers who share similar visible characteristics and attributes enhances the relevance of peer effects by emphasizing the more probable interpersonal relationships.

Value Homophily The remaining columns (8-10) present the estimated coefficients for the three spatial models based on the value homophily social network matrix, which, in addition to the status homophily matrix, assesses similarity based on academic performance, i.e., on the outcome variable of the students in each class.

Column (8) presents the coefficient estimates for the SAR model. The estimated endogenous peer effect becomes slightly larger in this specification. Compared to the LiM model (2), the estimated gender gap is smaller and closer to the OLS estimate, yet a bit larger than in the status homophily scenario (5). On average, boys outperform girls by 10.8738 points, a difference statistically significant at the 1% level. Like in the prior models, the direct effect (with feedback effects) is slightly larger, while the indirect effect is roughly twice its size. The adverse effect of student age is again reduced compared to all prior specifications, resulting in a local 10.7173 point decrease of performance if age is increased by one year and an 11.652 point decrease upon the inclusion of feedback mechanisms. The estimated regression coefficient and the direct effect of perceived bullying are also reduced in size. Interestingly however, the indirect effect increases slightly, i.e., the spillover effect for bullying gains importance. Although the changes in both the regression coefficients and the pseudo R^2 of 0.403 seem relatively small, the SAR model shows the largest value for the log-likelihood ($ll = -45,490$). Although no direct comparison between the models of identical parameterization is possible (i.e., between specifications (2),(5) and (7), respectively), the increase appears fairly large and suggestive of the advantage of this specification.

In column (9), once again the SLX model setup is tested. Comparable to the specifications (3) for LiM and (6) based on status homophily, the changes relative to the OLS baseline model in the estimated regression coefficients for individual level variables appear negligible. All of the coefficient estimates remain virtually unchanged in regard to sign, magnitude and significance. For the exogenous peer effects, the impact of peers using English as the language at home *almost always* increases slightly compared to specification (6), but remains higher than in the LiM case (3). Similarly, the adverse effect of peers being born outside the United States decreases in magnitude compared to the status homophily social network matrix, yet remains more pronounced than in the standard LiM case. The goodness of fit based on pseudo R^2 is identical to the other SLX specifications ($R_{pseudo}^2 = 0.333$), as well as the log-likelihood, which increases by one, as compared to specification (6).

The last model specification reported in column (10) shows the results for the SDM model with endogenous and exogenous peer effects. The endogenous peer effect is estimated to be 0.6957 and is statistically significant at the 1% level. Compared to the SDM model based on status homophily, the estimated gender gap increases slightly (coefficient and both direct and indirect effect). The adverse effect of student age in regard to coefficient and direct effect is reduced, however, the indirect effect increases in magnitude, indicating a higher negative spillover of age. The estimated regression coefficient and the direct effect of perceived bullying are negligibly smaller, while again, the indirect effect gains importance. Of the exogenous peer effects, only the negative externality of peers being born outside the United States remains significant, yet its magnitude is again reduced compared to the SLX scenario. In terms of pseudo R^2 , this model is identical to the LiM SDM. However, the log-likelihood of $ll = 45,486$ is considerably higher, indicating a better model choice.

While the relatively small differences in the specifications based on different social network matrices seem surprising at first, yet the overall susceptibility of spatial autoregressive models is often overstated, so that the small changes appear reasonable. (LeSage and Pace, 2011).

Robustness Analysis In the prior specifications in table 3, school environment factors have been explicitly included in the form of variables available from TIMSS. As a robustness analysis, the regressions are repeated, but this time with a dummy variable for each of the 213 schools in the sample, i.e., with school fixed effects, in order to verify that the results are not driven by unobserved school-level heterogeneity or unobserved common factors. These results are presented in table 5, while the overview of direct, indirect and total effects is shown in table 6.

Overall, the changes to the baseline results are small. For the individual characteristics, it is noteworthy that the effect of the individual speaking the language of the text at home sometimes becomes significant at the 5% level and negative in the SLX specifications (columns (3), (6) and (9)). In these specifications, as well as the baseline OLS model, the coefficient estimates for perceived bullying are markedly reduced, yet remain significant at the 1% level. It is noteworthy that considerable changes in the indirect effects of the student characteristics occur. While the adverse effect of peers being born outside the United States is lessened, the negative effect of peers using English at home only *sometimes* becomes significant at the 1% level, amounting to, e.g., a 95.579 point decrease of individual achievement in specification (9), in the case of all peers showing this characteristic.

In the SAR and SDM specifications, upon the inclusion of school fixed effects, the magnitude of the spatial autoregressive component, i.e., the endogenous peer effect, is reduced in size by approximately 0.10 points in each model. This finding is suggestive of a bias in the baseline specification, where parts of the school-level factors are falsely identified as positive peer effects instead. In all three SAR model specifications, the decreased endogenous effect entails a considerable reduce of the indirect effects, while the direct effects are altered only slightly. As an example, for the value homophily model (column (9)), the indirect effect of being a boy is decreased from 24.137 to 16.843 points upon the consideration of school fixed effects. The reduction of the direct effect from 11.822 to 11.774 points is almost negligible. The adverse effect of age, together with its direct effect gets even more pronounced, but still the indirect effect stays clearly below the magnitude of the baseline case. As a similar decrease in the indirect effects is visible for the SDM specifications, it can be concluded that the magnitude and thus the importance of spillover effects is reduced by including school fixed effects. However, the signs and significances of the effects remain unchanged, with the exception of the spillover of students being born outside the United States becoming insignificant.

Overall, including the school fixed effects increases the goodness of fit of the models, especially for the OLS and the SLX specifications, thus narrowing the gap between models with and without endogenous peer effects. However, the qualitative results of the prior chapter remain largely unchanged regarding signs and significances, which indicates that the baseline results are not driven by the omission of unobserved school level heterogeneity.

5 Conclusion

In this paper, I have shown that significant and positive peer effects are present for 8th grade students in secondary education in the United States after including a full set of individual-, parent- and school-level control variables. The application of spatial regression techniques enables an explicit modeling of peer relations and allows for a separation of direct and indirect effects of the respective variables to assess the magnitude of spillovers. It is worth noting that for various individual level factors, the indirect impact exceeds the magnitude of the direct effect. This strategy furthermore allows for a separate identification of the endogenous and exogenous peer effects described by [Manski \(1993\)](#), as well as contextual factors of the school environment. The peer relations within a class can be approximated by applying theoretical sociology-based approaches of homophily among members of the class as a closed social group. In the absence of information on actual relations between the nodes of the social network, different measures of similarity can be used to model the probability of interaction between all network members. This focus on more relevant peers strengthens the causal link between individual outcomes and the peer group influences and thus increases precision in determining peer groups.

The main findings have proven to be robust to changes in the theoretical foundation of the socio-economic spatial weights matrix, i.e., whether it is based on observable or non-observable characteristics, or it corresponds to an unweighted averaging of peer attributes. The inclusion of peer effects slightly increases the goodness of fit of the overall model and helps reducing omitted variable bias in the estimation of the remaining coefficients and furthermore renders school-level controls virtually obsolete. Additionally, the results are robust to the inclusion of school fixed effects rather than explicitly modeling them through variables provided in the TIMSS dataset.

The finding suggests that spatial regression approaches used for estimating peer effects are a suitable choice, consistent with recent findings and theoretical approaches ([Lee, 2006](#); [Lin, 2010](#); [Boucher et al., 2014](#)). Furthermore, relying on this estimation strategy circumvents the reflection problem discussed at length in the literature on peer effects in education. Describing social networks based on weak tie relations among students grouped in classes has proven to be a valid approach of modeling peer relations if no information on reported interactions are available. This increases the scope of potential countries suitable for studying peer effects and increases the credibility of causal impacts of peer performance. However, a validation using further country samples and possibly other datasets is yet to be done.

References

- Aesaert, K. and van Braak, J. (2015). Gender and socioeconomic related differences in performance based ict competences. *Computers & Education*, 84:8–25.
- Ammermueller, A. and Pischke, J.-S. (2009). Peer Effects in European Primary Schools: Evidence from the Progress in International Reading Literacy Study. *Journal of Labor Economics*, 27:315–348.
- Anger, S. and Heineck, G. (2010). Do Smart Parents Raise Smart Children? The Intergenerational Transmission of Cognitive Abilities. *Journal of Population Economics*, 23:1255–1282.
- Angrist, J. D. (2014). The perils of peer effects. *Labour Economics*, 30:98–108.
- Angrist, J. D. and Lang, K. (2004). Does school integration generate peer effects? evidence from boston’s metco program. *American Economic Review*, 94(5):1613–1634.
- Arcidiacono, P., Foster, G., Goodpaster, N., and Kinsler, J. (2012). Estimating spillovers using panel data, with an application to the classroom. *Quantitative Economics*, 3:421–470.
- Badland, H. and Schofield, G. (2006). Understanding the relationship between town size and physical activity levels: A population study. *Health & Place*, 12:538–546.
- Bedard, K. and Dhuey, E. (2006). The persistence of early childhood maturity: International evidence of long-run age effects. *Quarterly Journal of Economics*, 121(4):1437–1472.
- Bekiari, A., Pachi, V., and Hasanagas, N. (2017). Investigating bullying determinants and typologies with social network analysis. *Journal of Computer and Communication*, 5:11–27.
- Bekkerman, A. and Gilpin, G. (2013). High-speed internet growth and the demand for locally stable information content. *Journal of Urban Economics*, 77:1–10.
- Berry, C. R. and Glaeser, E. L. (2005). The divergence of human capital levels across cities. *Papers ins Regional Science*, 84:407–444.
- Bhatt, R. (2010). The impact of public library use on reading, television, and academic outcomes. *Journal of Urban Economics*, 68:148–166.
- Borgna, C. and Contini, D. (2014). Migrant achievement penalties in western europe: Do educational systems matter? *European Sociological Review*, 30(5):670–683.
- Boucher, V., Bramoullé, Y., Djebbari, H., and Fortin, B. (2014). Do peers affect student achievement? evidence from canada using group size variation. *Journal of Applied Econometrics*, 29:91–109.
- Bramoullé, Y., Djebbari, H., and Fortin, B. (2009). Identification of peer effects through social networks. *Journal of Econometrics*, 150:41–55.
- Brunello, G. and Rocco, L. (2013). The effect of immigration on the school performance of natives: Cross country evidence using pisa test scores. *Economics of Education Review*, 32:234–246.
- Burke, M. A. and Sass, T. R. (2013). Classroom peer effects and student achievement. *Journal of Labor Economics*, 31(1):51–82.

- Carrell, S. E., Fullerton, R. L., and West, J. E. (2009). Does your cohort matter? measuring peer effects in college achievement. *Journal of Labor Economics*, 27(3):439–464.
- Carrell, S. E. and Hoekstra, M. L. (2010). Externalities in the classroom: How children exposed to domestic violence affect everyone’s kids. *American Economic Journal: Applied Economics*, 2(1):211–228.
- Cebolla-Boada, H. (2011). Primary and secondary effects in the explanation of disadvantage in education: The children of immigrant families in france. *British Journal of Sociology of Education*, 32(3):407–430.
- Cöster, M. E., Fritz, J., Karlsson, C., Rosengren, B. E., and Karlsson, M. K. (2018). Extended physical education in children aged 6–15 years was associated with improved academic achievement in boys. *Acta Pædiatrica*, 107:1083–1087.
- Currie, J. (2009). Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development. *Journal of Economic Literature*, 47(1):87–122.
- Davis-Keane, P. E. (2005). The influence of parent education and family income on child achievement: The indirect role of parental expectations and the home environment. *Journal of Family Psychology*, 19(2):294–304.
- De Witte, K. and Geys, B. (2011). Evaluating efficient public good provision: Theory and evidence from a generalised conditional efficiency model for public libraries. *Journal of Urban Economics*, 69:319–327.
- Diette, T. M. and Uwaifo Oyelere, R. (2017). Gender and racial differences in peer effects of limited english students: A story of language or ethnicity? *IZA Journal of Migration*, 6(2):1–18.
- Epple, D. and Romano, R. E. (1998). Competition between private and public schools, vouchers, and peer-group effects. *American Economic Review*, 88(1):33–62.
- Fan, X. and Chen, M. (2001). Parental involvement and students’ academic achievement: A meta-analysis. *Educational Psychology Review*, 13(1):1–22.
- Farmer, T. W., Dawes, M., Hamm, J. V., Lee, D., Mehtaji, M., Hoffman, A. S., and Brooks, D. (2018). Classroom social dynamics management: Why the invisible hand of the teacher matters for special education. *Remedial and Special Education*, 39(3):177–192.
- Freeman, R. B., Machin, S., and Viarengo, M. (2010). Variation in educational outcomes and policies across countries and of schools within countries. *NBER Working Paper Series*, Working Paper 16293.
- Freeman, S., Theobald, R., Crowe, A. J., and Wenderoth, M. P. (2017). Likes attract: Students self-sort in a classroom by gender, demography, and academic characteristics. *Active Learning in Higher Education*, 18(2):115–126.
- Fryer, R. G. and Levitt, S. D. (2010). An Empirical Analysis of the Gender Gap in Mathematics. *American Economic Journal: Applied Economics*, 2(2):210–240.
- Gaviria, A. and Raphael, S. (2001). School-based peer effects and juvenile behavior. *Review of Economics and Statistics*, 83(2):257–268.
- Gower, J. C. (1971). A general coefficient of similarity and some of its properties. *Biometrics*, 27(4):857–871.
- Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, 78(6):1360–1380.

- Grygiel, P., Humenny, G., Rębisz, S., Bajcar, E., and Świtaj, P. (2018). Peer rejection and perceived quality of relations with schoolmates among children with ADHD. *Journal of Attention Disorders*, 22(8):738–751.
- Guiso, L., Monte, F., Sapienza, P., and Zingales, L. (2008). Culture, Gender, and Math. *Science*, 320:1164–1165.
- Hanushek, E. A., Kain, J. F., Markman, J. M., and Rivkin, S. G. (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics*, 18:527–544.
- Hanushek, E. A., Schwerdt, G., Wiederhold, S., and Woessmann, L. (2015). Returns to skills around the world: Evidence from PIAAC. *European Economic Review*, 73:103–130.
- Hanushek, E. A. and Woessmann, L. (2006). Does educational tracking affect performance and inequality? differences-in-differences evidence across countries. *Economic Journal*, 116:C63–C76.
- Hanushek, E. A. and Woessmann, L. (2011). The Economics of International Differences in Educational Achievement. In Hanushek, E. A., Machin, S., and Woessmann, L., editors, *Handbook of the Economics of Education. Band 3*, volume 26, pages 89–200. Elsevier.
- Harel-Fisch, Y., Walsh, S. D., Fogel-Grinvald, H., Amitai, G., Pickett, W., Molcho, M., Due, P., Gaspar de Matos, M., Craig, W., and Members of the HSBC Violence and Injury Prevention Focus Group (2011). Negative school perceptions and involvement in school bullying: A universal relationship across 40 countries. *Journal of Adolescence*, 34(4):639–652.
- Hughes, J. N., Cao, Q., West, S. G., Smith, P. A., and Cerda, C. (2017). Effect of retention in elementary grades on dropping out of school early. *Journal of School Psychology*, 65:11–27.
- Jackson, M., Jonsson, J. O., and Rudolphi, F. (2012). Ethnic inequality in choice-driven education systems: A longitudinal study of performance and choice in england and sweden. *Sociology of Education*, 85(2):158–178.
- Johnsen, A., Bjørknes, R., Iversen, A. C., and Sandbæk, M. (2018). School competence among adolescents in low-income families: Does parenting style matter? *Journal of Child and Family Structure*, 27:2285–2294.
- Kiss, D. (2017). How do ability peer effects operate? evidence on one transmission channel. *Education Economics*, 26(5):1–13.
- Knifsend, C. A., Camacho-Thompson, D. E., Juvonen, J., and Graham, S. (2018). Friends in activities, school-related affect, and academic outcomes in diverse middle schools. *Journal of Youth and Adolescence*, 47:1208–1220.
- Kretschmann, J., Vock, M., Lüdtke, O., and Gronostaj, A. (2016). Skipping to the bigger pond: Examining gender differences in students’ psychosocial development after early acceleration. *Contemporary Educational Psychology*, 46:195–207.
- Kristen, C., Reimer, D., and Kogan, I. (2008). Higher education entry of turkish immigrant youth in germany. *International Journal of Comparative Sociology*, 49(2–3):127–151.
- Lee, L.-F. (2006). Identification and estimation of econometric models with group interactions, contextual factors and fixed effects. *Journal of Econometrics*, 140:333–374.
- LeSage, J. P. and Pace, R. K. (2011). Pitfalls in higher order model extensions of basic spatial regression methodology. *The Review of Regional Studies*, 41(1):13–26.

- Lin, X. (2010). Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables. *Journal of Labor Economics*, 28(4):825–860.
- Louch, H. (2000). Personal network integration: Transitivity and homophily in strong-tie relations. *Social Networks*, 22:45–64.
- Lyle, D. S. (2007). Estimating and interpreting peer and role model effects from randomly assigned social groups at west point. *Review of Economics and Statistics*, 89(2):289–299.
- Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *Review of Economic Studies*, 60(3):531–542.
- Martins, L. and Veiga, P. (2010). Do inequalities in parents' education play an important role in PISA students' mathematics achievement test scores disparities? *Economics of Education Review*, 29:1016–1033.
- McCrary, J. and Royer, H. (2011). The effect of female education on fertility and infant health: Evidence from school entry policies using exact date of birth. *American Economic Review*, 101(1):158–195.
- McEwan, P. J. and Shapiro, J. S. (2007). The benefits of delayed primary school enrollment – discontinuity estimates using exact birth dates. *The Journal of Human Resources*, 43(1):1–28.
- McPherson, J. M. and Smith-Lovin, L. (1987). Homophily in voluntary organizations: Status distance and the composition of face-to-face groups. *American Sociological Review*, 52(3):370–379.
- McPherson, J. M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27:415–444.
- Moreland, R. L. and Beach, S. R. (1992). Exposure effects in the classroom: The development of affinity among students. *Journal of Experimental Social Psychology*, 28:255–276.
- Moretti, E. (2013). Real wage inequality. *American Economic Journal: Applied Economics*, 5:65–103.
- Nieuwenhuis, J. (2018). The interaction between school poverty and agreeableness in predicting educational achievement. *Personality and Individual Differences*, 127:85–88.
- Ponzo, M. (2013). Does bullying reduce educational achievement? an evaluation using matching estimators. *Journal of Policy Modeling*, 35:1057–1078.
- Resende Oliveira, F., Almeida de Menezes, T., Irfi, G., and Resende Oliveira, G. (2018). Bullying effect on student's performance. *Economia*, 19:57–73.
- Rindermann, H. (2007). The g-Factor of International Cognitive Ability Comparisons: The Homogeneity of Results in PISA, TIMSS, PIRLS and IQ-Tests Across Nations. *European Journal of Personality*, 21:667–706.
- Sacerdote, B. (2001). Peer effects with random assignment: Results for dartmouth roommates. *Quarterly Journal of Economics*, 116(2):681–704.
- Schuetz, G., Ursprung, H. W., and Woessmann, L. (2008). Education Policy and Equality of Opportunity. *Kyklos*, 61:279–308.

- Sorensen, L. C., Cook, P. J., and Dodge, K. A. (2017). From parents to peers: Trajectories in sources of academic influence grades 4 to 8. *Educational Evaluation and Policy Analysis*, 39(4):697–711.
- Steinberg, L. and Monahan, K. C. (2007). Age differences in resistance to peer influence. *Dev Psychol*, 43(6):1531–1543.
- Tasdemir, M. S. and Yalcin Arslan, F. (2018). Feedback preferences of EFL learners with respect to their learning styles. *Cogent Education*, 5(1):1–17.
- Todd, P. E. and Wolpin, K. I. (2003). On the Specification and Estimation of the Production Function for Cognitive Achievement. *The Economic Journal*, 113(485):F3–F33.
- Warne, R. T. (2017). Possible economic benefits of full-grade acceleration. *Journal of School Psychology*, 65:54–68.
- Warne, R. T. and Liu, J. K. (2017). Income differences among grade skippers and non-grade skippers across genders in the term sample, 1936–1976. *Learning and Instruction*, 47:1–12.
- Wechsler, H., Devereaux, R. S., Davis, M., and Collins, J. (2000). Using the school environment to promote physical activity and healthy eating. *Preventive Medicine*, 31:121–137.
- Winston, G. and Zimmerman, D. (2004). *Peer Effects in Higher Education*, pages 395–424. University of Chicago Press.

A Tables

Table 1: Summary Statistics for Categorical Variables

| Variable | Freq. | Percent | Cum. |
|---|-------|---------|--------|
| <i>Student gender</i> | | | |
| Girl* | 4,349 | 50.77 | 50.77 |
| Boy | 4,217 | 49.23 | 100.00 |
| <i>Student frequency of using test language at home</i> | | | |
| Always* | 6,334 | 73.94 | 73.94 |
| Almost always | 1,445 | 16.87 | 90.81 |
| Sometimes | 705 | 8.23 | 99.04 |
| Never | 82 | 0.96 | 100.00 |
| <i>Student born in United States</i> | | | |
| Yes* | 8,097 | 94.52 | 94.52 |
| No | 469 | 5.48 | 100.00 |
| <i>Highest education (Mother)</i> | | | |
| Some Primary or Lower secondary or did* | 312 | 3.64 | 3.64 |
| Lower secondary | 543 | 6.34 | 9.98 |
| Upper secondary | 1,792 | 20.92 | 30.90 |
| Short-cycle tertiary | 960 | 11.21 | 42.11 |
| Bachelor's or equivalent | 1,709 | 19.95 | 62.06 |
| Postgraduate degree | 1,166 | 13.61 | 75.67 |
| Don't know | 2,084 | 24.33 | 100.00 |
| <i>Highest education (Father)</i> | | | |
| Some Primary or Lower secondary or did* | 334 | 3.90 | 3.90 |
| Lower secondary | 592 | 6.91 | 10.81 |
| Upper secondary | 1,937 | 22.61 | 33.42 |
| Short-cycle tertiary | 782 | 9.13 | 42.55 |
| Bachelor's or equivalent | 1,320 | 15.41 | 57.96 |
| Postgraduate degree | 958 | 11.18 | 69.15 |
| Don't know | 2,643 | 30.85 | 100.00 |
| <i>Amount of books at home</i> | | | |
| 0–10 books* | 1,479 | 17.27 | 17.27 |
| 11–25 books | 1,862 | 21.74 | 39.00 |
| 26–100 books | 2,520 | 29.42 | 68.42 |
| 101–200 books | 1,420 | 16.58 | 85.00 |
| More than 200 | 1,285 | 15.00 | 100.00 |
| <i>Number of digital information devices</i> | | | |
| None* | 31 | 0.36 | 0.36 |
| 1–3 devices | 443 | 5.17 | 5.53 |
| 4–6 devices | 1,835 | 21.42 | 26.96 |
| 7–10 devices | 2,645 | 30.88 | 57.83 |
| More than 10 devices | 3,612 | 42.17 | 100.00 |
| <i>Number of home study supports</i> | | | |
| Neither Own Room nor Internet Connection | 150 | 1.75 | 1.75 |
| Either Own Room or Internet Connection | 1,695 | 19.79 | 21.54 |
| Both Own Room and Internet Connection | 6,721 | 78.46 | 100.00 |

table continued on next page

Table 1: Summary Statistics for Categorical Variables

| Variable | Freq. | Percent | Cum. |
|--|-------|---------|--------|
| <i>Mother born in US</i> | | | |
| Yes* | 6,569 | 76.69 | 76.69 |
| No | 1,729 | 20.18 | 96.87 |
| I don't know | 268 | 3.13 | 100.00 |
| <i>Father born in US</i> | | | |
| Yes* | 6,333 | 73.93 | 73.93 |
| No | 1,772 | 20.69 | 94.62 |
| I don't know | 461 | 5.38 | 100.00 |
| <i>How many people live in area?</i> | | | |
| More than 500,000* | 1,140 | 13.31 | 13.31 |
| 100,001 to 500,000 | 1,463 | 17.08 | 30.39 |
| 50,001 to 100,000 | 1,676 | 19.57 | 49.95 |
| 30,001 to 50,000 | 989 | 11.55 | 61.50 |
| 15,001 to 30,000 | 1,234 | 14.41 | 75.90 |
| 3,001 to 15,000 | 1,328 | 15.50 | 91.41 |
| 3,000 or fewer | 736 | 8.59 | 100.00 |
| <i>Immediate area of school location</i> | | | |
| Urban, densely populated* | 1,841 | 21.49 | 21.49 |
| Suburban, on fringe of urban area | 2,552 | 29.79 | 51.28 |
| Medium size city or large town | 2,145 | 25.04 | 76.33 |
| Small town or village | 1,603 | 18.71 | 95.04 |
| Remote rural | 425 | 4.96 | 100.00 |
| <i>Students from economically disadvantaged background</i> | | | |
| 0 to 10%* | 1,002 | 11.70 | 11.70 |
| 11 to 25% | 1,292 | 15.08 | 26.78 |
| 26 to 50% | 2,397 | 27.98 | 54.76 |
| More than 50% | 3,875 | 45.24 | 100.00 |
| <i>Students from economically affluent background</i> | | | |
| 0 to 10%* | 3,874 | 45.23 | 45.23 |
| 11 to 25% | 1,812 | 21.15 | 66.38 |
| 26 to 50% | 1,636 | 19.10 | 85.48 |
| More than 50% | 1,244 | 14.52 | 100.00 |
| <i>Percent of students that speak language of test</i> | | | |
| More than 90%* | 4,735 | 55.28 | 55.28 |
| 76 to 90% | 1,531 | 17.87 | 73.15 |
| 51 to 75% | 1,037 | 12.11 | 85.26 |
| 26 to 50% | 635 | 7.41 | 92.67 |
| 25% or less | 628 | 7.33 | 100.00 |

Notes: The summary statistics above correspond to the unweighted variables from the 8,566 student sample generated from the TIMSS 2015 release available from the IEA's website. Values that corresponded to missing at random or no answer provided have been removed as is suggested by the corresponding codebook.

* denotes the reference category for categorical variables.

Table 2: Summary Statistics for Metric Variables

| Variable | Mean | Std.Dev. | Min. | Max. |
|--|--------|----------|-------|--------|
| Student age | 14.235 | 0.467 | 9.830 | 18.330 |
| Perceived bullying | 10.009 | 1.936 | 2.467 | 12.784 |
| Parent home educational resources | 10.816 | 1.675 | 4.232 | 13.884 |
| School shortage of science resources | 10.815 | 1.816 | 3.886 | 14.859 |
| School shortage of mathematics resources | 10.778 | 1.730 | 4.050 | 14.554 |
| School reported discipline problems | 10.138 | 1.466 | 7.302 | 14.027 |
| School emphasis on academic success | 10.005 | 2.044 | 4.101 | 16.632 |

Notes: The summary statistics above correspond to the unweighted metric variables from the 8,566 student sample generated from the TIMSS 2015 release available from the IEA's website. Values that corresponded to missing at random or no answer provided have been removed as is suggested by the corresponding codebook.

Table 3: Baseline Regression Results for Linear-in-Means, Status Homophily and Value Homophily

| Variable | Adjacency Matrix | | | | | | | | | |
|--|-------------------------|-------------------------|--------------------------|-------------------------|-------------------------|--------------------------|--------------------------|-------------------------|--------------------------|-------------------------|
| | (1) OLS | (2) SAR | (3) SLX | (4) SDM | (5) SAR | (6) SLX | (7) SDM | (8) SAR | (9) SLX | (10) SDM |
| Student gender: Boy | 10.7848*** (1.3683) | 11.9304*** (1.0574) | 10.8240*** (1.3631) | 11.9311*** (1.0574) | 10.6090*** (1.0575) | 10.8796*** (1.3634) | 10.6230*** (1.0577) | 10.8738*** (1.0460) | 10.8779*** (1.3631) | 10.8897*** (1.0461) |
| Student age | -15.1701*** (1.4818) | -10.9742*** (1.1462) | -14.9393*** (1.4787) | -11.0120*** (1.1480) | -10.8748*** (1.1465) | -14.9146*** (1.4784) | -10.9017*** (1.1481) | -10.7173*** (1.1340) | -14.9162*** (1.4784) | -10.7439*** (1.1357) |
| Frequency of using language of test at home | | | | | | | | | | |
| Almost always | 1.4034 (2.2100) | 2.0700 (1.7077) | 1.3816 (2.2074) | 2.1005 (1.7122) | 2.1694 (1.7080) | 1.5249 (2.2078) | 2.2555 (1.7129) | 2.1211 (1.6894) | 1.5029 (2.2075) | 2.1912 (1.6942) |
| Sometimes | -3.2051 (3.0894) | -1.9581 (2.3872) | -3.3057 (3.0841) | -2.0139 (2.3923) | -1.9551 (2.3876) | -2.9992 (3.0839) | -1.8921 (2.3926) | -1.9743 (2.3616) | -3.0862 (3.0838) | -1.9411 (2.3668) |
| Never | -1.0493 (7.3161) | -1.3067 (5.6532) | -0.0498 (7.2906) | -1.2628 (5.6551) | -1.2106 (5.6542) | -0.7541 (7.2887) | -1.3156 (5.6547) | -1.3143 (5.5926) | -0.4771 (7.2885) | -1.3537 (5.5937) |
| Student born outside of United States | -7.7286** (3.2552) | -1.1950 (2.5166) | -5.4397* (3.2564) | -0.5082 (2.5266) | -0.8893 (2.5172) | 2.5001 (3.4967) | 2.2636 (2.7128) | -0.9285 (2.4897) | -0.1835 (3.3799) | 1.2684 (2.5940) |
| Perceived Bullying | 2.2639*** (0.3513) | 1.0247*** (0.2719) | 2.2587*** (0.3502) | 1.0269*** (0.2721) | 1.0259*** (0.2720) | 2.2646*** (0.3501) | 1.0297*** (0.2721) | 1.0019*** (0.2690) | 2.2626*** (0.3501) | 1.0046*** (0.2691) |
| Endogenous Peer Effect | | 0.6874*** (0.0085) | | 0.6855*** (0.0086) | 0.6865*** (0.0086) | | 0.6843*** (0.0086) | 0.6978*** (0.0084) | | 0.6957*** (0.0085) |
| Peer frequency of using language of test at home | | | | | | | | | | |
| Almost always | | | 29.9380*** (6.6725) | 1.7600 (5.1878) | | 25.4829*** (6.5328) | 0.8587 (5.0777) | | 27.5495*** (6.6050) | 1.3310 (5.0792) |
| Sometimes | | | -0.8667 (9.6110) | 6.7114 (7.4556) | | -2.9863 (9.5488) | 5.3141 (7.4088) | | -3.2366 (9.6250) | 5.3502 (7.3876) |
| Never | | | -5.3181 (27.2854) | 0.7423 (21.1647) | | -34.6935 (28.1655) | -10.8686 (21.8531) | | -25.3909 (28.3320) | -8.6560 (21.7448) |
| Peer born outside of United States | | | -80.0856*** (10.8790) | -24.4649*** (8.4674) | | -94.9039*** (12.8832) | -30.0047*** (10.0282) | | -92.8786*** (12.4636) | -27.6912*** (9.5982) |
| Parental Factors | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| School Factors | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 8,566 | 8,566 | 8,566 | 8,566 | 8,566 | 8,566 | 8,566 | 8,566 | 8,566 | 8,566 |
| Log Likelihood | -47556 | -45569 | -47521 | -45565 | -45575 | -47520 | -45570 | -45490 | -47519 | -45486 |
| Pseudo R ² | 0.327 | 0.402 | 0.333 | 0.406 | 0.401 | 0.333 | 0.405 | 0.403 | 0.333 | 0.406 |

Notes: Dependent variable is the average score in the TIMSS subjects of mathematics and science. Constant term is included, but not shown.

Parental Factors refer to controls for educational achievement, immigration background, and home economic and educational resources specific to the student's parents.

School Factors refer to controls for the school's location and economic environment, as well as variables describing the school's attitude towards academic success and reported discipline problems.

Robust standard errors are reported in parentheses.

*: Significant at the 10% level. **: Significant at the 5% level. ***: Significant at the 1% level.

Table 4: Overview of Direct, Indirect and Total Effects for Individual Level Student Characteristics

| Variable | Adjacency Matrix | | | | | | | | | |
|---|----------------------|-----------------------|-----------------------|------------------------|-----------------------|------------------------|------------------------|-----------------------|-------------------------|------------------------|
| | (1) OLS | (2) SAR | (3) SLX | (4) SDM | (5) SAR | (6) SLX | (7) SDM | (8) SAR | (9) SLX | (10) SDM |
| Student Gender: Boy | 10.785*** (1.368) | 12.901*** (1.145) | 10.824*** (1.363) | 12.89*** (1.144) | 11.48*** (1.145) | 10.88*** (1.363) | 11.484*** (1.144) | 11.822*** (1.138) | 10.878*** (1.363) | 11.828*** (1.137) |
| Student Age | -15.17*** (1.482) | -11.867*** (1.238) | -14.939*** (1.479) | -11.897*** (1.239) | -11.768*** (1.239) | -14.915*** (1.478) | -11.785*** (1.24) | -11.652*** (1.232) | -14.916*** (1.478) | -11.67*** (1.232) |
| Frequency of Using Language of Test at Home | | | | | | | | | | |
| Almost always | 1.403 (2.21) | 2.238 (1.847) | 1.382 (2.207) | 2.476 (1.92) | 2.348 (1.848) | 1.525 (2.208) | 2.54 (1.92) | 2.306 (1.837) | 1.503 (2.208) | 2.545 (1.917) |
| Sometimes | -3.205 (3.089) | -2.117 (2.581) | -3.306 (3.084) | -1.389 (2.693) | -2.116 (2.584) | -2.999 (3.084) | -1.416 (2.695) | -2.147 (2.568) | -3.086 (3.084) | -1.446 (2.691) |
| Never | -1.049 (7.316) | -1.413 (6.113) | -0.5 (7.291) | -1.277 (6.603) | -1.31 (6.118) | -0.754 (7.2890) | -2.71 (6.652) | -1.429 (6.081) | -0.477 (7.2890) | -2.542 (6.6610) |
| Student born outside of United States | -7.729** (3.255) | -1.292 (2.721) | -5.44 (3.256) | -3.419 (2.817) | -0.962 (2.724) | 2.5 (3.497) | -1.107 (2.73) | -1.009 (2.707) | -1.184 (3.38) | -2.052 (2.737) |
| Perceived Bullying | 2.264*** (.351) | 1.108*** (.294) | 2.259*** (.35) | 1.109*** (.294) | 1.11*** (.294) | 2.265*** (.35) | 1.113*** (.294) | 1.089*** (.292) | 2.263*** (.35) | 1.091*** (.292) |
| Student Gender: Boy | | 25.251*** (2.461) | | 25.026*** (2.44) | 22.343*** (2.392) | | 22.154*** (2.37) | 24.137*** (2.51) | | 23.941*** (2.487) |
| Student Age | | -23.227*** (2.553) | | -23.098*** (2.538) | -22.903*** (2.537) | | -22.735*** (2.519) | -23.79*** (2.644) | | -23.62*** (2.627) |
| Frequency of Using Language of Test at Home | | | | | | | | | | |
| Almost always | | 4.381 (3.619) | 29.917*** (6.668) | 9.791 (16.093) | 4.569 (3.603) | 25.465*** (6.5280) | 7.321 (15.704) | 4.708 (3.756) | 27.53*** (6.6000) | 9.023 (16.293) |
| Sometimes | | -4.144 (5.054) | -866 (9.6040) | 16.312 (23.157) | -4.118 (5.03) | -2.984*** (9.542) | 12.248 (22.915) | -4.382 (5.244) | -3.234 (9.618) | 12.64 (23.703) |
| Never | | -2.766 (11.966) | -5.314 (27.266) | -377 (65.8820) | -2.549 (11.908) | -34.669 (28.146) | -35.864 (67.688) | -2.917 (12.415) | -25.373 (28.312) | -30.33 (69.878) |
| Student born outside of United States | | -2.529 (5.324) | -80.03*** (10.871) | -75.926*** (25.885) | -1.873 (5.299) | -94.837*** (12.874) | -86.714*** (28.865) | -2.061 (5.524) | -92.8140*** (12.455) | -84.718*** (29.166) |
| Perceived Bullying | | 2.169*** (.577) | | 2.154*** (.5720) | 2.161*** (.5740) | | 2.147*** (.5690) | 2.224*** (.599) | | 2.209*** (.593) |

table continued on next page

Table 4: Overview of Direct, Indirect and Total Effects for Individual Level Student Characteristics –continued–

| Variable | Adjacency Matrix | | | Status Homophily | | | Value Homophily | | | |
|---|----------------------|-----------------------|-------------------------|------------------------|-----------------------|------------------------|------------------------|-----------------------|------------------------|-----------------------|
| | (1) OLS | (2) SAR | (3) SLX | (4) SDM | (5) SAR | (6) SLX | (7) SDM | (8) SAR | (9) SLX | (10) SDM |
| Student Gender: Boy | 10.785*** (1.368) | 38.152*** (3.552) | 10.824*** (1.363) | 37.916*** (3.53) | 33.823*** (3.494) | 10.88*** (1.363) | 33.638*** (3.47) | 35.96*** (3.602) | 10.878*** (1.363) | 35.769*** (3.577) |
| Student Age | -15.17*** (1.482) | -35.094*** (3.746) | -14.939*** (1.479) | -34.995*** (3.732) | -34.671*** (3.733) | -14.915*** (1.478) | -34.52*** (3.715) | -35.442*** (3.832) | -14.916*** (1.478) | -35.29*** (3.815) |
| Frequency of Using Language of Test at Home | | | | Total Effects | | | | | | |
| Almost always | 1.403 (2.21) | 6.619 (5.465) | 31.299*** (6.926) | 12.267 (17.096) | 6.916 (5.45) | 26.99*** (6.7940) | 9.8610 (16.71) | 7.0150 (5.591) | 29.033*** (6.863) | 11.568 (17.322) |
| Sometimes | -3.205 (3.089) | -6.2620 (7.635) | -4.172 (9.961) | 14.924 (24.576) | -6.2330 (7.613) | -5.9830 (9.8990) | 10.832 (24.339) | -6.529 (7.811) | -6.321 (9.972) | 11.194 (25.159) |
| Never | -1.049 (7.316) | -4.179 (18.079) | -5.364 (28.25) | -1.655 (69.6680) | -3.859 (18.027) | -35.423 (29.113) | -38.574 (71.5520) | -4.346 (18.495) | -25.85 (29.278) | -32.872 (73.839) |
| Student born outside of United States | -7.729*** (3.255) | -3.822 (8.045) | -85.4690*** (11.062) | -79.345*** (27.281) | -2.835 (8.023) | -92.337*** (12.038) | -87.821*** (29.587) | -3.07 (8.231) | -92.997*** (11.966) | -86.77*** (30.181) |
| Perceived Bullying | 2.264*** (.351) | 3.277*** (.869) | 2.259*** (.35) | 3.263*** (.864) | 3.271*** (.867) | 2.265*** (.35) | 3.26*** (.861) | 3.313*** (.889) | 2.263*** (.35) | 3.3*** (.884) |

Notes: Partitioned effects are calculated based on the estimated regression coefficients presented in table 3 (see section B.1 of the appendix for details). The dependent variable is the average score in the TIMSS subjects of mathematics and science.
Robust standard errors are reported in parentheses.

*: Significant at the 10% level. **: Significant at the 5% level. ***: Significant at the 1% level.

Table 5: Robustness Analysis: School Fixed Effects

| Variable | Adjacency Matrix | | | | | | | | | |
|--|-------------------------|-------------------------|--------------------------|--------------------------|-------------------------|--------------------------|--------------------------|-------------------------|--------------------------|--------------------------|
| | (1) OLS | (2) SAR | (3) SLX | (4) SDM | (5) SAR | (6) SLX | (7) SDM | (8) SAR | (9) SLX | (10) SDM |
| Student Gender: Boy | 11.0863*** (1.2210) | 12.1091*** (1.0597) | 11.3760*** (1.2143) | 12.2032*** (1.0596) | 10.9649*** (1.0595) | 11.1976*** (1.2158) | 10.9850*** (1.0601) | 11.1939*** (1.0492) | 11.2787*** (1.2149) | 11.2483*** (1.0496) |
| Student Age | -15.7301*** (1.3640) | -12.7210*** (1.1850) | -15.4371*** (1.3564) | -12.6664*** (1.1846) | -12.5970*** (1.1851) | -15.3808*** (1.3573) | -12.5138*** (1.1847) | -12.4183*** (1.1736) | -15.3691*** (1.3567) | -12.3383*** (1.1733) |
| Frequency of Using Language of Test at Home | | | | | | | | | | |
| Almost always | 0.9933 (1.9783) | 1.7258 (1.7168) | 1.2351 (1.9754) | 1.6120 (1.7235) | 1.8236 (1.7168) | 1.2062 (1.9754) | 1.7234 (1.7224) | 1.7923 (1.7001) | 1.2657 (1.9750) | 1.7033 (1.7063) |
| Sometimes | -4.3314 (2.7729) | -2.8688 (2.4065) | -7.0545** (2.7796) | -3.8604 (2.4260) | -2.8276 (2.4064) | -6.3616** (2.7773) | -3.5653 (2.4223) | -2.8234 (2.3830) | -6.6373** (2.7774) | -3.6328 (2.4001) |
| Never | -1.3356 (6.5384) | -1.4250 (5.6740) | -2.5305 (6.5668) | -1.9558 (5.7296) | -1.2437 (5.6739) | -4.5642 (6.5594) | -2.6074 (5.7196) | -1.3557 (5.6188) | -4.0345 (6.5614) | -2.5279 (5.6687) |
| Student born outside of United States | -2.2917 (2.9230) | 0.1920 (2.5370) | -3.1271 (2.9121) | -0.0678 (2.5415) | 0.4729 (2.5370) | 5.6007* (3.1102) | 3.4885 (2.7123) | 0.4242 (2.5123) | 2.9725 (2.9856) | 2.3098 (2.5794) |
| Perceived Bullying | 1.6750*** (0.3163) | 0.9832*** (0.2748) | 1.6545*** (0.3145) | 0.9865*** (0.2747) | 0.9869*** (0.2748) | 1.6603*** (0.3146) | 0.9899*** (0.2746) | 0.9635*** (0.2721) | 1.6547*** (0.3145) | 0.9654*** (0.2720) |
| Endogenous Peer Effect | | 0.5933*** (0.0109) | | 0.5876*** (0.0111) | 0.5918*** (0.0109) | | 0.5864*** (0.0111) | 0.6090*** (0.0108) | | 0.6035*** (0.0109) |
| Peer frequency of Using Language of Test at Home | | | | | | | | | | |
| Almost always | | | 3.5288 (9.1030) | -8.4721 (7.9457) | | -3.8144 (8.7023) | -10.3977 (7.5890) | | -0.4869 (8.8859) | -9.1763 (7.6785) |
| Sometimes | | | -99.7045*** (13.6443) | -36.1397*** (11.9652) | | -88.8534*** (13.2311) | -32.5034*** (11.5859) | | -95.6463*** (13.5069) | -33.7497*** (11.7224) |
| Never | | | -53.4310* (31.7302) | -20.3125 (27.6919) | | -97.0201*** (31.5797) | -40.2579 (27.5568) | | -86.0673*** (32.1756) | -36.4436 (27.8120) |
| Peer born outside of United States | | | -62.6472*** (13.4709) | -21.0462* (11.7797) | | -67.8233*** (15.2125) | -24.6433* (13.2896) | | -71.6723*** (15.1266) | -23.5559* (13.0972) |
| Parental Factors | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| School Factors | No | No | No | No | No | No | No | No | No | No |
| School Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 8,566 | 8,566 | 8,566 | 8,566 | 8,566 | 8,566 | 8,566 | 8,566 | 8,566 | 8,566 |
| Log Likelihood | -46462 | -45392 | -46411 | -45383 | -45394 | -46416 | -45385 | -45321 | -46412 | -45312 |
| Pseudo R ² | 0.479 | 0.512 | 0.485 | 0.513 | 0.511 | 0.484 | 0.514 | 0.513 | 0.485 | 0.515 |

Notes: Dependent variable is the average score in the TIMSS subjects of mathematics and science. Constant term is included, but not shown. Robust standard errors are reported in parentheses.
 *, Significant at the 10% level. **, Significant at the 5% level. ***, Significant at the 1% level.

Table 6: Overview of Direct, Indirect and Total Effects for Individual Level Student Characteristics based on School Fixed Effects

| Variable | Adjacency Matrix | | | Status Homophily | | | Value Homophily | | | |
|---|----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| | (1) OLS | (2) SAR | (3) SLX | (4) SDM | (5) SAR | (6) SLX | (7) SDM | (8) SAR | (9) SLX | (10) SDM |
| Student Gender: Boy | 11.086*** (1.221) | 12.677*** (1.111) | 11.376*** (1.214) | 12.757*** (1.109) | 11.484*** (1.11) | 11.198*** (1.216) | 11.49*** (1.109) | 11.774*** (1.104) | 11.279*** (1.215) | 11.813*** (1.103) |
| Student Age | -15.73*** (1.364) | -13.318*** (1.239) | -15.437*** (1.356) | -13.241*** (1.237) | -13.194*** (1.24) | -15.381*** (1.357) | -13.089*** (1.238) | -13.062*** (1.233) | -15.369*** (1.357) | -12.958*** (1.231) |
| Frequency of Using Language of Test at Home | | | | | | | | | | |
| Almost always | .993 (1.978) | 1.807 (1.797) | 1.235 (1.975) | 1.031 (1.957) | 1.91 (1.798) | 1.206 (1.975) | .988 (1.945) | 1.885 (1.788) | 1.266 (1.975) | 1.025 (1.955) |
| Sometimes | -4.331 (2.773) | -3.003 (2.519) | -7.054 (2.78) | -6.8260 (2.806) | -2.962 (2.52) | -6.362 (2.777) | -6.276 (2.788) | -2.97 (2.506) | -6.6370 (2.777) | -6.623 (2.808) |
| Never | -1.336 (6.538) | -1.492 (5.94) | -2.53 (6.567) | -3.613 (6.6400) | -1.303 (5.9430) | -4.564 (6.559) | -5.8820 (6.619) | -1.426 (5.91) | -4.035 (6.561) | -5.687 (6.674) |
| Student born outside of United States | -2.292 (2.923) | .201 (2.656) | -3.127 (2.912) | -1.696 (2.86) | .495 (2.657) | 5.601* (3.11) | 1.718 (2.681) | .446 (2.643) | 2.973 (2.986) | .466 (2.702) |
| Perceived Bullying | 1.675*** (.316) | 1.029*** (.288) | 1.655*** (.314) | 1.031*** (.287) | 1.034*** (.288) | 1.66*** (.315) | 1.035*** (.287) | 1.013*** (.286) | 1.655*** (.314) | 1.014*** (.285) |
| Student Gender: Boy | | 17.088*** (1.692) | | 16.819*** (1.657) | 15.364*** (1.635) | | 15.057*** (1.602) | 16.843*** (1.749) | | 16.547*** (1.714) |
| Student Age | | -17.952*** (1.821) | | -17.458*** (1.782) | -17.651*** (1.803) | | -17.153*** (1.766) | -18.086*** (1.915) | | -18.15*** (1.874) |
| Frequency of Using Language of Test at Home | | | | | | | | | | |
| Almost always | | 2.435 (2.426) | 3.526 (9.097) | -17.651 (19.024) | 2.555 (2.409) | -3.812 (8.696) | -21.943 (18.112) | 2.697 (2.562) | -487 (8.8800) | -19.86 (19.119) |
| Sometimes | | -4.048 (3.399) | -99.6350 (13.635) | -90.095 (28.55) | -3.962 (3.374) | -88.791 (13.222) | -80.868 (27.559) | -4.248 (3.589) | -95.5790 (13.497) | -87.604 (29.084) |
| Never | | -2.011 (8.0080) | -53.394 (31.708) | -50.343 (66.527) | -1.743 (7.9510) | -96.952 (31.558) | -97.682 (65.8450) | -2.04 (8.455) | -86.0070 (32.153) | -92.544 (69.381) |
| Student born outside of United States | | .271 (3.58) | -62.603 (13.461) | -49.462 (28.012) | .663 (3.556) | -67.776 (15.202) | -52.825 (29.994) | .638 (3.781) | -71.622 (15.116) | -54.016 (31.296) |
| Perceived Bullying | | 1.387*** (.39) | 1.36*** (.381) | 1.36*** (.381) | 1.383*** (.387) | 1.357*** (.379) | 1.357*** (.379) | 1.45*** (.412) | 1.42*** (.402) | 1.42*** (.402) |

table continued on next page

Table 6: Overview of Direct, Indirect and Total Effects for Individual Level Student Characteristics based on School Fixed Effects –continued–

| Variable | Adjacency Matrix | | | Status Homophily | | | Value Homophily | | | |
|---|----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|----------------------|-----------------------|-----------------------|-----------------------|
| | (1) OLS | (2) SAR | (3) SLX | (4) SDM | (5) SAR | (6) SLX | (7) SDM | (8) SAR | (9) SLX | (10) SDM |
| Student Gender: Boy | 11.086*** (1.221) | 29.765*** (2.739) | 11.376*** (1.214) | 29.576*** (2.7) | 26.848*** (2.691) | 11.198*** (1.216) | 26.547*** (2.656) | 28.617*** (2.797) | 11.279*** (1.215) | 28.36*** (2.758) |
| Student Age | -15.73*** (1.364) | -31.269*** (2.994) | -15.437*** (1.356) | -30.699*** (2.952) | -30.845*** (2.978) | -15.381*** (1.357) | -30.242*** (2.94) | -31.747*** (3.083) | -15.369*** (1.357) | -31.108*** (3.041) |
| Frequency of Using Language of Test at Home | | | | Total Effects | | | | | | |
| Almost always | .993 (1.978) | 4.242 (4.222) | 4.761 (9.487) | -16.621 (20.094) | 4.465 (4.206) | -2.606 (9.081) | -20.956 (19.162) | 4.582 (4.349) | .779 (9.267) | -18.835 (20.213) |
| Sometimes | -4.331 (2.773) | -7.0520 (5.916) | -106.689 (14.26) | -96.9210 (30.17) | -6.924 (5.893) | -95.1530 (13.826) | -87.1440 (29.148) | -7.218 (6.093) | -102.217 (14.107) | -94.227 (30.741) |
| Never | -1.336 (6.538) | -3.503 (13.948) | -55.924 (33.284) | -53.956 (70.4120) | -3.045 (13.893) | -101.516 (33.058) | -103.563 (69.687) | -3.466 (14.365) | -90.042 (33.679) | -98.2310 (73.389) |
| Student born outside of United States | -2.292 (2.923) | .472 (6.236) | -65.73 (13.947) | -51.158 (29.512) | 1.158 (6.213) | -62.175 (14.478) | -51.107 (30.526) | 1.084 (6.423) | -68.6500 (14.79) | -53.55 (32.235) |
| Perceived Bullying | 1.675*** (.316) | 2.417*** (.6760) | 1.655*** (.314) | 2.391*** (.666) | 2.416*** (.673) | 1.66*** (.315) | 2.392*** (.664) | 2.463*** (.6960) | 1.655*** (.314) | 2.434*** (.6860) |

Notes: Partitioned effects are calculated based on the estimated regression coefficients presented in table 3 (see section B.1 of the appendix for details). Dependent variable is the average score in the TIMSS subjects of mathematics and science. Robust standard errors are reported in parentheses.
 *: Significant at the 10% level. **: Significant at the 5% level. ***: Significant at the 1% level.

Figures

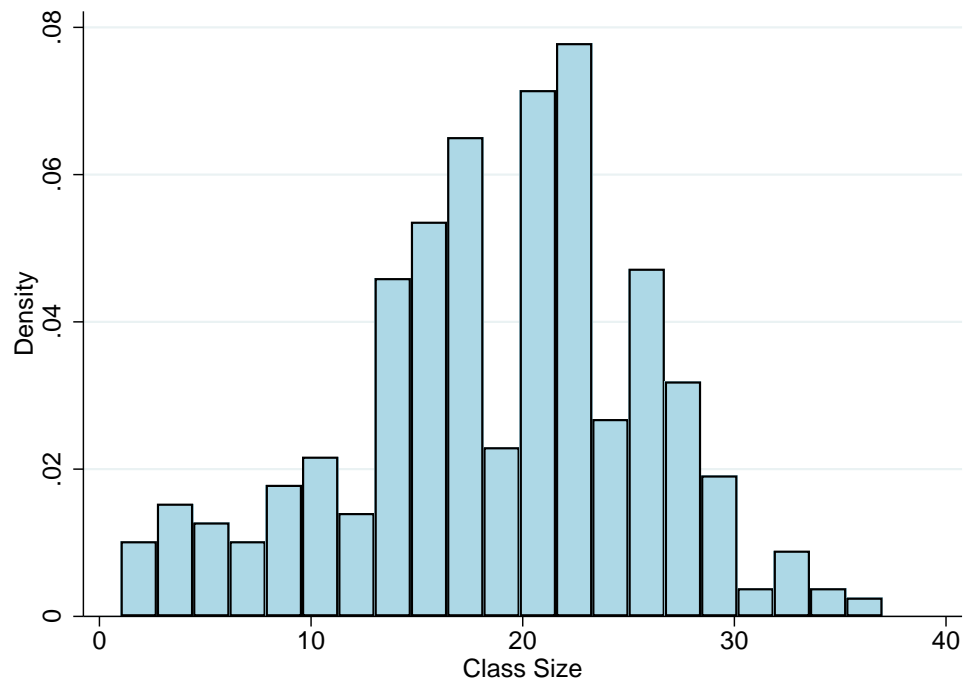


Figure 1: Histogram of Class Sizes in the Estimation Sample

B Technical Appendix

B.1 Interpreting Parameter Estimates

In spatial regression models, coefficient estimates cannot be simply interpreted as marginal effects, due to spillover or feedback mechanisms. Rather, effects of changes in the explanatory variables can be partitioned into direct, indirect and total effects, which differ from the estimated regression coefficients. The derivation of the partitioned effects is illustrated using an example based on the spatial autoregressive model specification.

For an SAR model, a simplified version of the regression equation (11), solved for O , can be portrayed as equation (17), where \mathbf{Z} is a matrix of all K variables¹⁰ used in the prior regression and θ is the corresponding coefficient vector.

$$O = (\mathbf{I}_n - \rho \mathbf{W})^{-1} [\alpha \mathbf{1}_n + \mathbf{Z}\theta + \varepsilon] \quad (17)$$

Refraining from the matrix notation for $\mathbf{Z}\theta$, the equation can be rearranged and rewritten so that z_r is now a $n \times 1$ vector for the observations of variable $r \in [1, \dots, K]$.

$$O = V(\mathbf{W}) [\alpha \mathbf{1}_n + \varepsilon] + \sum_{r=1}^K S_r(\mathbf{W}) z_r \quad (18)$$

with $S_r(\mathbf{W}) = V(\mathbf{W})(\mathbf{I}_n \theta_r)$
and $V(\mathbf{W}) = (\mathbf{I}_n - \rho \mathbf{W})^{-1}$

Expanding the matrices $S_r(\mathbf{W})$, \mathbf{O} and vector z_r yields the following equation.

$$\begin{pmatrix} o_1 \\ o_2 \\ \vdots \\ o_n \end{pmatrix} = \sum_{r=1}^K \begin{bmatrix} S_r(\mathbf{W}_{11}) & S_r(\mathbf{W}_{12}) & \cdots & S_r(\mathbf{W}_{1n}) \\ S_r(\mathbf{W}_{21}) & S_r(\mathbf{W}_{22}) & \cdots & S_r(\mathbf{W}_{2n}) \\ \vdots & \vdots & \ddots & \vdots \\ S_r(\mathbf{W}_{n1}) & S_r(\mathbf{W}_{n2}) & \cdots & S_r(\mathbf{W}_{nn}) \end{bmatrix} \begin{pmatrix} z_{1r} \\ z_{2r} \\ \vdots \\ z_{nr} \end{pmatrix} + V(\mathbf{W}) [\alpha \mathbf{1}_n + \mathbf{W}\mathbf{X}\lambda + \varepsilon] \quad (19)$$

For a single observation, e.g., student i , this can be rewritten as

$$o_i = \sum_{r=1}^K [S_r(\mathbf{W})_{i1} z_{1r} + S_r(\mathbf{W})_{i2} z_{2r} + \cdots + S_r(\mathbf{W})_{in} z_{nr}] + V(\mathbf{W})_i [\alpha + \varepsilon] \quad (20)$$

Taking the derivative with respect to variable r of individual j then yields the following expression

$$\frac{\partial o_i}{\partial z_{jr}} = S_r(\mathbf{W})_{ij} \quad (21)$$

For variable r , the effect of a change of individual $j \in [1, \dots, n]$ on individual $i \in [1, \dots, n]$ can be portrayed in a $n \times n$ matrix of derivatives. For $i \neq j$, this matrix illustrates that, different than in the standard OLS case, the derivative of o_i with respect to z_{jr} is non-zero. Changes in characteristics of individual j affect not only j 's outcome, but also the outcomes of all other individuals.

¹⁰This includes student, parent and school level variables

$$S_r(\mathbf{W}) = \begin{bmatrix} \frac{\partial o_1}{\partial z_{1r}} & \frac{\partial o_1}{\partial z_{2r}} & \cdots & \frac{\partial o_1}{\partial z_{nr}} \\ \frac{\partial o_2}{\partial z_{1r}} & \frac{\partial o_2}{\partial z_{2r}} & \cdots & \frac{\partial o_2}{\partial z_{nr}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial o_n}{\partial z_{1r}} & \frac{\partial o_n}{\partial z_{2r}} & \cdots & \frac{\partial o_n}{\partial z_{nr}} \end{bmatrix} = \begin{bmatrix} S_r(\mathbf{W})_{11} & S_r(\mathbf{W})_{12} & \cdots & S_r(\mathbf{W})_{1n} \\ S_r(\mathbf{W})_{21} & S_r(\mathbf{W})_{22} & \cdots & S_r(\mathbf{W})_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ S_r(\mathbf{W})_{n1} & S_r(\mathbf{W})_{n2} & \cdots & S_r(\mathbf{W})_{nn} \end{bmatrix} \quad (22)$$

From the matrix in equation (22), the partitioned effects can be read in the following way. First, the *direct effects* can be found on the main diagonal of the matrix. The average direct effect corresponds to the average of all elements on the main diagonal, i.e.,

$$\overline{direct}_r = n^{-1} \text{tr}(S_r(\mathbf{W})) \quad (23)$$

The *total effect* on individual i resulting from a one-unit change in variable z_r for all individuals (including i) can be read from the sum of the i th row of the matrix. Accordingly, the average total effect corresponds to the average of the sum of all rows, which can be expressed as in equation (24).

$$\overline{total}_r = n^{-1} \iota'_n \text{tr}(S_r(\mathbf{W})) \iota_n \quad (24)$$

Last, the *indirect effect* on the outcome of student i corresponds to a change in variable z_r for all other individuals, except for i . This corresponds to the difference between the individual total and direct effect. The average indirect effect then is the average total minus the average direct effect.

$$\overline{indirect}_r = \overline{total}_r - \overline{direct}_r \quad (25)$$

The results presented in the tables 4 and 6 have been constructed according to the formulas introduced above, using the built-in post-estimation command *estat impact* from the *sp* class commands introduced in Stata version 15. By default, this command uses the delta method to calculate the variance of the impacts. Individual achievement is observed for individual i in class c of school s . The school draws its students from neighborhood n in town t of country l , so achievement is O_{icsntl} .

A potential experiment would be moving an individual into a different class or school. Or into a different neighborhood or town.