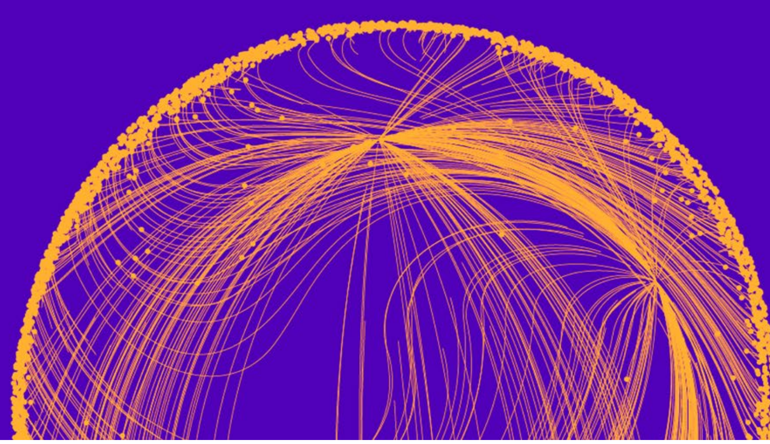
HELSINKI GSE DISCUSSION PAPERS 37 · 2025

Effects of Disruptive Peers in Endogenous Social Networks

Torsten Santavirta
Miguel Sarzosa



Helsinki GSE Discussion Papers

Helsinki GSE Discussion Papers 37 · 2025

Torsten Santavirta and Miguel Sarzosa:
Effects of Disruptive Peers in Endogenous Social Networks

ISBN 978-952-7543-36-8 (PDF)
ISSN 2954-1492

Helsinki GSE Discussion Papers:
https://www.helsinkigse.fi/discussion-papers

Helsinki Graduate School of Economics
PO BOX 21210
FI-00076 AALTO
FINLAND

Helsinki, May 2025

Effects of Disruptive Peers in Endogenous Social Networks*

Torsten Santavirta† Miguel Sarzosa‡

May 3, 2025

Abstract

This study uses sociometric data to show that social connections in the classroom shape the diffusion of the negative externalities on cognitive achievement generated by abused and neglected peers. We find the strongest negative effects for students who are socially closest to the abused and neglected peer. The fade-out rate of the negative externality is such that being three peers away from an abused and neglected peer is equivalent to having no such peers. Although the inverse effect-distance relation applies to both verbal and numeric ability, it is conferred through different mechanisms. The abused and neglected peer's lower verbal ability harms her friends' verbal ability, whereas it is the disruptiveness itself that harms classmates' numeric ability.

JEL Classification: I21, J13, J24

*This paper was supported by the Handelsbanken Research Foundation grant number P2016-0093 and the Swedish Research Council, Grant No. 2017-00878. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Handelsbanken Research Foundation or the Swedish Research Council. We thank Ylva Almquist and Jannike Kjellström at the Centre for Health Equity Studies (CHESS) for their help with the data construction. We are grateful to Eric Auerbach, Manuel Bagues, Tim Bond, Vincent Boucher, Emily Breza, Eric Chyn, Gordon Dahl, Magne Mogstad, Kevin Mumford, Matthew Lindquist, Áureo de Paula, Michele Pellizzari, Victoria Prowse, Eirini Tatsi, and participants in seminars at Purdue Econ (West Lafayette, 2017), SOFI (Stockholm, 2018), HECER (Helsinki, 2018), SOLE (Toronto, 2018), University of York (York, UK, 2018), Purdue HDFS (West Lafayette, 2018), Ohio State University (Columbus, 2018), LACEA (Guayaquil, Ecuador, 2018), 3rd Applied Micro Workshop (Bolzano, Italy, 2019), 5th Conference on Network Science and Economics (IU, Bloomington, 2019), and AEA-ASSA (2021) for useful comments.

†Department of Economics, University of Helsinki. Helsinki, Finland; Institute for Housing and Urban Research, Uppsala University, Uppsala, Sweden; and Swedish Institute for Social Research, Stockholm University, Stockholm, Sweden. Email: torsten.santavirta@ibf.uu.se

‡Corresponding Author. Department of Economics, Purdue University. 403 W. State St, West Lafayette, IN 47907. Email: msarzosa@purdue.edu. Phone: (765)-494-5517

1 Introduction

Disruptive students have been shown to cause negative externalities on peers' achievement scores in elementary school (Aizer, 2008; Carrell and Hoekstra, 2010). The effects persist into the early career labor market (Carrell et al., 2018). Yet, beyond their existence and total magnitude, little is known about these important externalities. In particular, studies that examine peer effects as complex mechanisms governing the diffusion of the treatment effects in the classroom—and the resulting treatment heterogeneity—are scarce (Jackson and Yariv, 2011; Sacerdote, 2011; Carrell et al., 2013; Arduini et al., 2015; de Paula, 2017).[1] This area of research has, however, considerable potential to help guide policymakers towards interventions that can dissipate and contain the spread of the externality (Manski, 2013a,b; Graham, 2018).

Advances in the econometric modeling of peer effects using network data have opened up important avenues for research into how they permeate the reference group. Recent network versions of the Manski (1993) linear-in-means peer effects model (LIMM) provide ways to jointly estimate a network formation technology and a production function of economic outcomes whose inputs include peers' actions and characteristics, and prior similarities between agents (Goldsmith-Pinkham and Imbens, 2013; Arduini et al., 2015; Hsieh and Lee, 2015; Hsieh et al., 2020; Johnsson and Moon, 2021; Auerbach, 2022a,b; Hsieh et al., 2022). Yet, few applications have exploited these models in educational settings and, to our knowledge, none of the existing applications deal with negative externalities.

This paper leverages network data to explore the diffusion of negative externalities generated by abused and neglected classmates. Specifically, we ask whether the size of the externality on cognitive ability wanes with social distance between the student and her abused and neglected peer. Our focus is thus on the interaction-driven heterogeneity

[1] In fact, in his handbook chapter on peer effects, Sacerdote (2011) observes that the "linear-in-means model masks considerable heterogeneity in the effects experienced by different types of students."

of treatment responses induced by the way the network architecture shapes the diffusion of a treatment. Starting from the network version of the LIMM, we derive and estimate a model in which distance (the shortest path length in terms of links in the network) between the abused and neglected peer and each student in the classroom affects the size of the externality. We use data from the Stockholm Birth Cohort Study—a unique longitudinal dataset following the entire cohort of children born in Stockholm, Sweden in 1953—that links school survey data with administrative records including birth records and population censuses. These data have two important features. First, we can identify all children who have been abused or neglected by their parents, as recorded in the universe of administrative files of the Swedish child protection services agency. This feature is crucial for causal identification as exposure to parental abuse and neglect is a good proxy for student disruptiveness (see Table A.1 in the Appendix and Carrell and Hoekstra (2010) and Carrell et al. (2018)), while also being exogenous to the abused and neglected student's classmates.[2] Second, the data include classroom friendship nominations in the sixth grade, making it possible to map the social network of each classroom and, in particular, the position of the abused and neglected peers within it. Meanwhile, cognitive tests administered in sixth grade measure verbal, numeric, and spatial ability.

Identification in our empirical strategy relies on within-school across-classroom variation of the fraction of abused and neglected classmates (Hoxby, 2000).[3] We show empirically that disruptive peers within the same school were randomly assigned to classrooms. In

[2]Existing literature in psychology and economics shows that child maltreatment significantly contributes to various social and emotional problems, such as aggression, depression, anxiety, and decreased social competence (Carlson, 2000; Wolfe et al., 2003; Holt et al., 2008; Moylan et al., 2009; Carrell and Hoekstra, 2010; Eriksen et al., 2014; Sarzosa and Urzua, 2021). Table A.1 shows empirically that the children in our data exposed to parental abuse and neglect have substantially lower cognitive test scores and grades and are more likely to have adjustment problems and engage in risky behaviors compared to their counterparts who were not abused or neglected.

[3]Earlier studies have used variations of this approach to quantify negative classroom externalities exerted by children, with particular attention to those with attention deficit hyperactivity disorders (Aizer, 2008; de Chaisemartin and Navarrete H., 2023), migrants (Gould et al., 2009), children linked to domestic violence (Carrell and Hoekstra, 2010; Carrell et al., 2018), boys with female-sounding names (Figlio, 2005), students with high blood lead levels (Gazze et al., 2022), low-ability students (Lavy et al., 2011), and bullies (Sarzosa and Urzua, 2021; Sarzosa, 2021).

this sense, we exploit the fact that a student in a given school may be allocated by chance to a classroom with a low number of abused and neglected peers, while another who attends the same school and grade may be allocated to the classroom next door where there is a high number of abused and neglected peers. In our case, the strategy is facilitated by the institutional framework that administered school assignment in Sweden in the 1960s. In particular, students attended their nearest school, and tracking on the basis of ability or background was not permitted, resulting in randomly formed classroom rosters within elementary schools, at least in terms of the share of abused and neglected students.

Causal identification also requires dealing with the fact that social networks are endogenous, as students themselves choose who to befriend and who not to befriend. To this end, we model the friendship link formation process as the product of homophilic preferences (i.e., people tend to befriend others who are similar to themselves) and unobserved student-level social ability (Graham, 2017). We measure pair-wise homophily using information on arguably exogenous sociodemographic characteristics such as father's socioeconomic status (SES) at birth, whether the student is the first-born sibling, gender, prenatal care, block-level geographic location of residence, and other predetermined data. The results of the friendship formation model allow us to generate instruments for the observed friendship links.

Our findings indicate that abused and neglected peers exert the largest negative effect on the verbal and numeric abilities of their closest friends. The externality fades out at a rate proportional to the distance. When the distance between the student and the abused and neglected peer exceeds two friendship links in the social network, the statistically significant negative peer effect fades out completely. This pattern is particularly evident for males, whereas for girls we only find evidence of effect heterogeneity by social distance for numeric ability. We further exploit the classroom networks to estimate the structural parameters of the LIMM using the joint regression framework (network formation equation and outcome equation) developed by Johnsson and Moon

(2021), which controls for the independent effect of an individual's social ability on academic achievement. This allows us to distinguish between different sources of treatment transmission. In particular, we can discern how much of the negative externality is generated by variation in the child abuse and neglect of peers (exogenous effects, in the LIMM jargon) and how much is due to spillovers through the peers' academic achievement (endogenous effects). We find that abused and neglected peers affect their classmates' verbal ability scores through the endogenous effect. In contrast, the effect of abused and neglected peers on their classmates' numeric ability seems to be more directly afflicted by the abuse and neglect itself. We use Monte Carlo simulations to interpret our structural estimates. The structural model results corroborate the decay of the peer effect that we find in our social distance model.

We contribute to the growing literature on school peer effects by documenting that social networks determine their scope and diffusion. Unlike existing studies on school peer effects, we show that an abused or neglected student can have heterogeneous consequences for her peers.[4] These consequences depend on how far apart the student and the abused and neglected peer are in the classroom network. In this regard, our paper relates to the literature that views peer effects as a mechanism by which the treatments may propagate through the network (Calvó-Armengol et al., 2009; Jackson and Yariv, 2011; Banerjee et al., 2013; Arduini et al., 2015; Schennach, 2018) and the literature that puts forth theoretical and empirical models of peer effects based on social networks incorporating the idea that the position in the network matters for exposure (Jackson and Wolinsky, 1996; Calvó-Armengol et al., 2009; Bramoullé et al., 2009; Dahl et al., 2014; Díaz et al., 2021). In particular, we extend the literature focusing on the role of networks in the diffusion, reach, and interaction-driven heterogeneity of the effects of treatments and interventions in educational settings (List et al., 2020; Opper, 2019). To the best of our knowledge, we are the first to show that the externality of disrup-

[4]Patacchini et al. (2017) provide some evidence on effect heterogeneity by friendship duration, observing that strong friendships—those that last for more than a year—have a greater scope of influencing behaviors.

tive students is almost exclusively confined to their circle of friends and these friends' friends, and that the mechanisms through which the effects propagate and decay in the numeric-ability dimension differ from those through which they propagate in the verbal-ability dimension. Our contribution is important because, as corroborated by our counterfactual policy exercises, it indicates that policies aimed at reducing the negative consequences of disruptive peers should target disruptive children and their circle of friends. Accordingly, class-wide interventions are not the most cost-efficient way of addressing this problem given that some resources would be devoted to unaffected students. Examples of potentially effective targeted local interventions include providing special emotional support to children with abusive or neglectful parents, and introducing remedial language courses for disruptive male students and their friends. Finally, our results suggest that teachers and education authorities, if provided with pertinent information, may be able to anticipate which students are more likely to belong to the disruptive peer's social circle and thus prepare targeted early interventions.

The remainder of the paper is organized as follows. Section 2 describes the data. Section 3 presents the empirical strategies for estimating both our reduced-form and diffusion models of peer effects and discusses the validity of the underlying identifying assumptions. Section 4 shares the main estimates of disruptive peer effects on students' cognitive outcomes. Section 5 provides evidence of the mechanisms driving the negative effects of disruptive peers by outlining and estimating the structural LIMM. Section 6 tests the robustness of our results and Section 7 concludes.

2 Data: The Stockholm Birth Cohort Study

We use the original sampling frame of the Stockholm Birth Cohort Study (hereafter, SBC) of all children born in the Stockholm metropolitan area in 1953 (Almquist, 2014).[5]

[5]See Stenberg and Vågerö (2006) for a cohort profile and Santavirta and Sarzosa (2025) for codebooks for the included registries.

In total, 12,677 children completed a school survey in sixth grade in 1966. They comprise our baseline sample of interest. We drop classrooms with fewer than seven students as they presumably hold very different social networks to ones observed in regular-sized classrooms. We also drop 54 schools with only one class per grade and which therefore are not subject to the classroom-level variation within school that our empirical strategy exploits, and 30 schools with at least one special education class in sixth grade.[6] We end up with a study sample of 7,995 students from 382 classrooms belonging to 116 schools. Table 1 presents summary statistics of the key variables used in this study.

An advantage of the SBC's 1966 school survey is that it was conducted in-class, which made practically every sixth grader in the county of Stockholm fill out two questionnaires (during two consecutive classes), eliciting information on a number of learning-related aspects including friendship nominations and a cognitive ability test. As a result, the non-response rate is only 7 percent (the percentage of students absent on that particular school day). The low non-response rate in combination with the fact that all students in the county took the survey is likely to increase the external validity of our study.[7]

The SBC also contains the municipal registers of case files of investigations by the child protection services (known as the Child Welfare Committee, hereafter CWC). These registers record each child's cases accumulated up to age 18 and case files remain in the register even though families would move to other municipalities.

Moreover, extensive information on the cohort members and their families was ascertained. Among other things, SBC contains prenatal and perinatal care records, student

[6]We use the direct individual-level survey responses on the attendance of a special education classroom to identify schools with special education classes. In our data, there are 435 students reported being in a sixth-grade special education class. Using cluster analyses, we further validate that these special education classes house children whose IQ scores are substantially lower than scores of students in the same schools but in different classrooms.

[7]In addition to the 12,677 sixth graders who completed the survey, another 1,296 cohort members completed the survey but were in fifth grade of seventh grade at the time of the survey in 1966 due to grade retention of class skipping. These students are dropped from our data. Our reported non-response rate refers to all 14,073 students who completed the survey as separate non-response rates were not available at grade-level.

performance and Census records containing data such as household composition, parents' education and occupation and the family's neighborhood block of residence at the time the cohort member was born. Importantly for our purposes, the data contain information on all cohort members in the Stockholm region (net of the 7 percent non-response rate) allowing us with the help of the friendship nominations of the school survey to characterize the social networks of complete classrooms. Below, we describe the key components of the data in detail. For further information on the data used in this study, see Santavirta and Sarzosa (2025)

Cognitive tests. We use the three components of ability collected by the school survey in 1966 (at age 13): numeric, verbal, and spatial ability. The tests were developed by the Swedish Institute for Educational Research in the early 1960s and have served to this date as the default cognitive tests in elementary school (Svensson, 1964). The test of numeric ability poses 40 numerical sequences of six numbers, each of which follows a logical pattern based on elementary arithmetic concepts. The students are asked to predict the next two numbers following the same pattern in the sequence. The verbal ability test presents the student with 40 words, for which the student has to find antonyms among four options. The spatial ability test consists of 40 unfolded figures that need to be folded mentally.

The verbal and numeric tests are weighted more toward crystallized intelligence. Scores on crystallized intelligence tests are in part determined by innate ability but also by acquired skills and knowledge and are thus depending on educational opportunity and motivation (Borghans et al., 2008). Some work on ability testing suggests that the numeric and verbal tests might more appropriately be called achievement tests than intelligence test (Almlund et al., 2011). In contrast, the spatial ability test is weighted more towards fluid intelligence, which is often considered the more innate of the two measures of intelligence, cleaner from acquired knowledge (Svensson, 1971). In addition to the cognitive test scores, we use the grade point average (GPA) of the Spring term

Table 1: Summary Statistics

Variable	Mean	SD	Obs.
School characteristics			
Number of schools			116
Number of classrooms			382
Avg no of classrooms/school	3.29	1.36	116
Students per classroom	25.96	3.62	382
Abused and neglected peers (A&N)			
Share of classrooms with a A&N	0.628	0.484	382
No of A&N in the classroom	1.168	1.311	382
No of A&N —(A&N in classroom$\geq$1)	1.858	1.205	240
Overall share of A&N	0.056	0.230	7,995
Controls			
Female†	0.512	0.500	7,995
Parents social aid recipients†	0.118	0.322	7,995
Birthweight (g)	3518.3	529.2	6,476
Mother's age	24.21	5.69	6,655
Owner of dwelling†	0.181	0.385	7,995
Dwelling size†	0.916	0.277	7,995
Older siblings	0.860	0.994	7,995
Outcomes			
Verbal ability	25.698	6.088	7,995
Numeric ability	21.752	7.589	7,995
Spatial ability	23.493	6.808	7,995
GPA in grade 9 (in hundredths))	321.473	76.462	7,592

Note: The summary statistics for disruptive peers within classrooms as well as the individual-level descriptive statistics (outcomes and controls) refer to the analytic sample (see Section 2 for sample restrictions). The second column reports standard deviations (SD). † Indicates a binary variable.

in grade nine of compulsory school. This grade point average is measured on a grade scale from 1 to 5 (we rescale the variable multiplying it by 100).

Social interactions. In a classroom survey conducted for sixth graders (age 13), students were asked to nominate their three best classroom friends (the nominations only concerned friends within the same classroom). Of all students in SBC who participated in the school survey and nominated friends (n=11,854), 7,499 nominated three friends (62.7 percent), 3,198 nominated two friends (26.8 percent), 787 nominated only one friend (6.6 percent), and 468 did not nominate any friends (3.9 percent). In our analytic sample (n=7,995), 5,124 students nominated three friends (64.1 percent), 2,120 students nominated two friends (26.5 percent), 488 students nominated only one friend (6.1 percent), and 262 students did not nominate any friends (3.2 percent). Reassuringly, our study sample is representative of the complete population of survey participants in terms of social interactions. A limitation to the sociometric data is that roughly 7 percent of cohort members did not participate in the school survey, yet the students who participated in the survey could have made nominations to these classmates who were absent on the day of the survey. We do not observe these nominations. Moreover, we do not observe the nominations made to classmates in sixth grade who were not born in 1953 and hence did not belong to the studied birth cohort. Based on the share of the cohort members themselves who deviated from the default grade at the time of the completion of the survey (sixth grade), this source of attrition is estimated to be 9 percent. Compared to the missing values in other youth surveys ascertaining youth friendship networks, this share of missing values seems reasonably low (Boucher and Houndetoungan, 2023).[8]

Disruptiveness. We proxy disruptive peers with children whose parents have undergone an investigation for parental abuse or neglect up until the child was 13 years old (in sixth grade) by the CWC. The Swedish child welfare policy was very institutionalized already in the 1960s with a well established system of foster care mandated by the Child Protection Act. The share of foster care placements in the population of underaged children at a given point in time was roughly the same in the 1960s as in

[8]This type of attrition, though relatively small, can potentially be selective if retained students (born in 1952) are more likely to be abused and neglected. We find that retained students from the cohort we observe (born in 1953) are indeed 7 percentage points more likely to be abused and neglected than cohort members being in the right grade for their age. In Section 4 and Section VII in the Web Appendix we show that our results remain robust to the inclusion of the attrited retainers.

the contemporary context, i.e., 1 percent of the relevant population (SOU1974:7, 1974; Statistics Sweden, 2023). The CWC investigated parents on suspicion of abuse and neglect and 90 percent of the investigations in our data were substantiated and lead to protective actions by the CWC (warnings, intensified surveillance or child removal). We drop the unsubstantiated cases. Of all children in SBC (n=12,677), we observed 798 CWC investigations that match our definition of abuse and neglect, of which 689 (86.3 percent) were carried out before the child started school and the rest were carried out during elementary school before age 13.[9] In our study sample (n=7,995), we observe 446 children whose parents were investigated for abuse and neglect at least once before the child turned 13. Of these, 385 (86.3 percent) children's parents were investigated before the child started school. Of all the 446 investigations in our analytic sample, 401 cases warranted child removal and placement in foster care or institution. According to our data, the main reasons for launching an investigation for abuse and neglect were parents' alcohol abuse, parents' psychiatric disorders, and parents' death.

3 Empirical model and identification

3.1 Social distance and the diffusion of disruptive peer effects

In this section, we lay out a framework that builds on the network version of the (Manski, 1993) LIMM formalized in Bramoullé et al. (2009). We explore whether the externality that a particular student exerts on each of her classmates may differ depending on how closely socially connected the two of them are. We develop an empirical model

[9] Of the 798 observed investigations in the full SBC sample that match our definition for abuse and neglect, 651 children were removed from their biological families mandated by §31 of the Swedish Child Protection Act (in co-operation with the parents) and eight child removals were mandated by §29 of the same act (against the will of the biological parents). The rest (n=139) of the cases led to preventive measures (warnings, instructions, advice or supervision) based on §26.

of how social influence diffuses with social distance, define our estimating equations and outline an identification strategy based on dyadic link formation.

3.1.1 Linear-in-means model of social interactions

Children interact more with some classmates than with others. Thus, the graph that represents a typical classroom's social network differs from a uniform and complete set of connections between all those who belong to the group (Coleman, 1964; de Paula, 2017). Formally, we consider classroom r to be a set of students $C_r = \{1, \dots, n_r\}$ that create a graph $(n_r, \mathbf{D}_r)$ through friendships. That is, a network $\mathcal{C}_r$ with vertex set C_r and link set $D_r = \{(i, j) \in C_r : i \text{ considers } j \text{ to be her friend}\}$. Therefore, we can represent $\mathcal{C}_r$ with the adjacency matrix $\mathbf{D}$ whose typical element $D_{ij} = 1$ when i considers j to be her friend and 0 otherwise. Following convention, we assume that $D_{ii} = 0$, $\forall i$. That is, one cannot befriend oneself (Jackson, 2010). Importantly for our setting, within classroom r, there is a subset of students $A_r = \{i \in C_r : i\text{'s parents were investigated by CWC for abuse and neglect}\}$.

Student i in classroom r is described by: (i) whether the her parents at some point during their childhood were investigated by CWC for abuse and neglect $a_{ir} \in \{0, 1\}$, (ii) her scholastic achievement y_{ir}, which is the student's choice variable, and (iii) her position in classroom network $\mathbf{D}_r$. The student wants to do well in school but achievement is costly and is influenced by peers. In particular, students want to comply with social norms, defined as the average achievement of her friends: $m_{ir}^{-1}\mathbf{d}_{i,r}\mathbf{y}_r$, where $\mathbf{d}_{i,r}$ is $\mathbf{D}_r$'s i^{th} row and m_{ir} collects i's total number of friends $m_i = \sum_{j=1}^{n_r} D_{ij}$, such that vector $\mathbf{M}_r$ contains the network's degree sequence. If we define the row-normalized matrix $\mathbf{G} = \text{diag}(\mathbf{M})^{-1}\mathbf{D}$, we can write the student i's utility in a linear-quadratic structure similar to the one used in Topa and Zenou (2014).

$$U(y_{ir}, \mathbf{y}_r, \mathbf{G}_r) = (\alpha_0 + \alpha a_{ir} + \alpha_{\bar{x}}\mathbf{G}_r\mathbf{a}_r + \tilde{\eta}_r)\, y_{ir} - \frac{1}{2}y_{ir}^2 + \frac{\beta_{\bar{y}}}{2}(y_{ir} - \mathbf{G}_r\mathbf{y}_r)^2, \qquad (1)$$

where $\alpha_0 > 0$, $\alpha < 0$, $\alpha_{\bar{x}} < 0$, and $0 < \beta_{\bar{y}} < 1$. Disruptiveness affects student i's utility through two channels: it affects the marginal utility of i's effort ($\alpha_{\bar{x}}\mathbf{G}_r\mathbf{a}_r y_{ir}$) but also affects utility through the student's conformity to the peer's normative studying effort ($\frac{\beta_{\bar{y}}}{2}(y_{ir} - \mathbf{G}_r\mathbf{y}_r)^2$). Students' social sub-utility is such that the studying effort is positively or negatively affected by the degree to which they conform to their peers' level of studying effort, as is reflected by the last term of equation (1). This is in line with the findings of Liu et al. (2014) and Boucher et al. (2024), who both show that, for studying effort, students tend to conform to the social norm of their friends. The lower bound for $\beta_{\bar{y}}$ is based on the assumption that students have positive utility from conforming to their friends.

Stacking students into classroom vectors $\mathbf{y}_r$ and $\mathbf{a}_r$, each of size $n_r \times 1$, we can write the best-response function calculating the first order condition with respect to y_{ir}:

$$\mathbf{y}_r = \beta_0\iota_r + \beta\mathbf{a}_r + \beta_{\bar{y}}\mathbf{G}_r\mathbf{y}_r + \beta_{\bar{x}}\mathbf{G}_r\mathbf{a}_r + \eta_r, \tag{2}$$

where $\beta_0 = (1+\beta_{\bar{y}})^{-1}\alpha_0$, $\beta = (1+\beta_{\bar{y}})^{-1}\alpha$, $\beta_{\bar{x}} = (1+\beta_{\bar{y}})^{-1}\alpha_{\bar{x}}$ and $\eta_r = (1+\beta_{\bar{y}})^{-1}\tilde{\eta}_r$. Equation (2) is a version of the LIMM (Manski, 1993; Brock and Durlauf, 2001; Blume et al., 2011) that incorporates the social interactions collected in network $\mathcal{C}_r$ (Bramoullé et al., 2009; Goldsmith-Pinkham and Imbens, 2013). Let $\mathbf{I}_r$ be the $n_r \times n_r$ identity matrix, then Patacchini and Zenou (2012) and Ushchev and Zenou (2020) show that as long as $\beta_{\bar{y}} < 1$, which makes the matrix $(I_r - \beta_{\bar{y}}\mathbf{G}_r)$ row-diagonally dominant due to the row-normalized nature of $\mathbf{G}_r$, and thus invertible, the best-response functions produce a unique interior Nash equilibrium $\mathbf{y}^*$ given by

$$\mathbf{y}_r^* = (\mathbf{I}_r - \beta_{\bar{y}}\mathbf{G}_r)^{-1}(\beta_0\iota_r + \beta\mathbf{a}_r + \beta_{\bar{x}}\mathbf{G}_r\mathbf{a}_r + \eta_r), \tag{3}$$

which yields a reduced-form relation between peer abuse and neglect and achievement.

3.1.2 The diffusion of peer effects

As in Calvó-Armengol et al. (2009), let $\mathbf{W_u} = (\mathbf{I} - \beta_{\bar{y}}\mathbf{G}^{-1})\mathbf{u} = \sum_{k=0}^{\infty} \beta_{\bar{y}}^{k}\mathbf{G}^{k}\mathbf{u}$ be the vector of $\mathbf{u}$-weighted centralities of parameter $\beta_{\bar{y}}$ and network $\mathbf{G}$. Letting $\mathbf{u} = \beta_{\bar{x}}\mathbf{G}_r\mathbf{a}_r$, we can re-parameterize and write equation (3) in terms of a vector of centralities

$$\mathbf{y}_r^* = \phi_0\iota_r + \phi_1\mathbf{a}_r + \beta_{\bar{x}}\mathbf{W}_{\mathbf{G}_r\mathbf{a}_r}(\mathbf{G}, \beta_{\bar{y}}) + \varepsilon_r, \tag{4}$$

where $\phi_0 = (\mathbf{I}_r - \beta_{\bar{y}}\mathbf{G}_r)^{-1}\beta_0$, $\phi_1 = (\mathbf{I}_r - \beta_{\bar{y}}\mathbf{G}_r)^{-1}\beta$, and $\varepsilon = (\mathbf{I}_r - \beta_{\bar{y}}\mathbf{G}_r)^{-1}\eta_r$. Note that

$$\mathbf{b}(\mathbf{G}_r, \beta_{\bar{y}}, \mathbf{a}_r) = \mathbf{W}_{\beta_{\bar{y}}\mathbf{G}_r\mathbf{a}_r}(\mathbf{G}, \beta_{\bar{y}}) = \beta_{\bar{y}}\mathbf{G}_r\mathbf{a}_r + \beta_{\bar{y}}^2\mathbf{G}_r^2\mathbf{a}_r + \beta_{\bar{y}}^3\mathbf{G}_r^3\mathbf{a}_r + \dots$$

is the vector of Katz-Bonacich centralities, but taking into account only paths that lead to the abused and neglected (A&N) peers. That is, while $\beta_{\bar{y}}$ captures how the value of being connected to another node decays with distance, $\beta_{\bar{y}}\mathbf{a}_r$ captures the base value of each node. Given that a_{ir} $\forall i \in C_r$ is a binary variable, the latter implies that the only valuable nodes here are those who belong to set A_r (i.e., the A&N peers). Then, we can write equation (4) as

$$\mathbf{y}_r^* = \phi_0\iota_r + \phi_1\mathbf{a}_r + \frac{\beta_{\bar{x}}}{\beta_{\bar{y}}}\mathbf{b}(\mathbf{G}_r, \beta_{\bar{y}}, \mathbf{a}_r) + \varepsilon_r \tag{5}$$

This simplifies matters because the Katz-Bonacich centrality takes into account the distance between nodes as a factor in the calculation of node centrality. The centrality score of a node is proportional to the sum of the centrality scores of its neighbors, and so on recursively, until all nodes in the network have been accounted for. This recursive calculation of centrality takes into account the paths of different lengths between nodes. Nodes with high centrality tend to be connected to other highly central nodes, and these connections form pathways that can reduce the distance between nodes in the network. In contrast, nodes with low centrality tend to be connected to other low-centrality nodes, which can lead to longer geodesic distances between nodes in the network. Thus, we have a parametric relation between peers' disruptiveness $\mathbf{a}_r$ and one's own outcome $\mathbf{y}_r$, the strength of which is determined by social distance.

We can further write the centrality of node i as follows:

$$b_{ir} = \sum_{k=1}^{\infty} \sum_{j=1}^{n_r} \beta_{\bar{y}}^k g_{ij}^k a_{jr} = \sum_{j \in A_r} \sum_{k=\delta_{ij}}^{\infty} \beta_{\bar{y}}^k g_{ij}^k = \sum_{j \in A_r} \sum_{k=0}^{\infty} \beta_{\bar{y}}^{\delta_{ij}+k} g_{ij}^{\delta_{ij}+k},$$

where $\delta_{ij} > 0$ is the geodesic between nodes i and j, where $i \neq j$ and $\delta_{ii} = 0$. It indicates the minimum number of friendship links separating student i from student j in the social network. The second equality reflects the way we have defined the centrality in which the relevant paths are only those that lead to students in set A_r. It follows from the fact that each relevant node contributes to i's centrality depending on its distance to i at every path that connects these two nodes. That is, j's contribution to i's centrality is the weighted sum of all the paths that connect them, where the weight is proportional to the distance between i and j in each path. Therefore, there must exist one path among these in which i and j are closest. That is the geodesic between i and j, which is thus the path through which j contributes the most to i's centrality. The third equality takes the argument further as it sorts the contributions of j to i's centrality according to how long the paths that connect them are. Contributions to i's centrality through longer paths are more heavily discounted. How much an A&N peer can affect i' centrality depends on how close the two of them are socially. If we split each node's contribution to i's centrality made through the shortest path(s) from the rest of the contributions, we can write i's centrality as:

$$b_{ir}(\delta_{i\cdot}, g_{i\cdot}) = \sum_{j \in A_r} \beta_{\bar{y}}^{\delta_{ij}} g_{ij}^{\delta_{ij}} + \sum_{j \in A_r} \sum_{k=\delta_{ij}+1}^{\infty} \beta_{\bar{y}}^k g_{ij}^k. \quad (6)$$

Closer A&N nodes will influence i's centrality more because their greatest contribution will be discounted by less ($\beta_{\bar{y}}^{\delta_{ij}}$ will be larger), and will tend to have more paths connecting them to i than further away A&N nodes. That is, $\partial b_{ir}(\delta_{ij}, g_{ij})/\partial \delta_{ij} < 0$

$\forall j \in A_r$. We use this feature of centrality and the fact that empirical analyses of social networks have established that the diameter (the maximum distance between any pair of agents) of a social network tends to be small (Jackson and Rogers, 2007), in order to conceptualize a parsimonious model for the mechanisms through which effects of disruptive peers play out within the classroom. We *approximate* $b_{ir}(\delta_{ij})$ with a function of the geodesic between the two relevant nodes $\delta_{ij} a_{ij}$. Let $\mathbf{\Delta}_r$ be a symmetric matrix containing the geodesics δ_{ij} between any two nodes in set C_r. Then, we can write an approximation of (5) as

$$\mathbf{y}_r = \phi_0 \iota_r + \phi_1 \mathbf{a}_r + \phi_2 \mathbf{f}(\mathbf{\Delta}_r)\mathbf{a}_r + \hat{\boldsymbol{\varepsilon}}_r. \tag{7}$$

Note that $\mathbf{f}(\mathbf{\Delta}_r)$ captures variation coming exclusively from the first part of (6). Thus, it is capturing a lower bound of the effect that A&N students can have on their peers, as it does not include part of the effect that can be channeled through longer paths. Equation (7) is informative because $\nabla \mathbf{y}_{ir}/\nabla \mathbf{a}_r = \phi_1 \mathbf{e} + \phi_2 \mathbf{f}(\boldsymbol{\delta}_{ir}) + \nabla \hat{\boldsymbol{\varepsilon}}_{ir}/\nabla \mathbf{a}_r$, where $\mathbf{e}$ is a vector of zeroes with a one in the i^{th} position. It indicates that the size of the total peer-effect can depend on the distance between the two agents (i.e., the shortest path length between them).

In principle, we do not need to impose further restrictions on $\mathbf{f}(\mathbf{\Delta}_r)$ beyond the one bounding $\beta_{\bar{y}}$ to the $(0, 1)$ interval. However, when taking the function to our data, we need to make some considerations. First, we must consider its endogeneity, which arises from that of $\mathbf{D}$. Therefore, we need to ensure $\mathbf{f}(\mathbf{\Delta}_r)$ fulfills the rank condition with respect to the number of available instruments. Second, our data has two characteristics that we must accommodate through our choice of the functional form of $\mathbf{f}(\mathbf{\Delta}_r)$. The first characteristic is that due to the fact that social networks are built based on students' active choices to befriend some peers but not others, we observe several classes' social networks comprising two or more *components* (i.e., a subnetwork formed by path-connected nodes that is not connected with other subnetworks) resulting in students who, although being classmates, are not socially connected. We cannot calculate the shortest path length between two students belonging to two different components. That implies that the matrices $\mathbf{\Delta}_r$ in classrooms with disjoint subnetworks have some elements for which $\delta_{ij} = \infty$. In response to this feature of the data, we structure $\mathbf{f}(\mathbf{\Delta}_r)$ in a way in which we split it into two parts. One that captures possible effects of those who are further than κ links away (i.e., $\delta_{ij} > \kappa$, those with whom one shares a classroom but are not in one's component), and another one capturing the effect-heterogeneity by distance to peers with whom the student shares a component. That is, $\mathbf{f}(\mathbf{\Delta}_r) = \mathbf{f_1}(\mathbf{1}[\mathbf{\Delta}_r \leq \kappa]) + \mathbf{f_2}(\mathbf{1}[\mathbf{\Delta}_r > \kappa])$, where $\mathbf{1}[\cdot]$ is an indicator function that takes the value one if the argument is true and zero otherwise.

The second characteristic of our data that influences our choice of $\mathbf{f}(\mathbf{\Delta}_r)\mathbf{a}_r$ is the fact that some classrooms have no disruptive children. That is, those classrooms' $\mathbf{a}_r$ is a vector of zeros (i.e., $\iota_r \mathbf{a}_r = 0$, where ι_r is a row-vector of ones). We use those classrooms as a comparison group that helps us tease out classroom-wide effects from localized ones of disruptive peers. If social connections drive the diffusion of the effect, the outcomes of the disruptive students' classmates who are not socially connected to them should not differ from those of students who have no disruptive peers in their classroom. If, on the contrary, just having a disruptive peer in the classroom is enough to disrupt learning regardless of social connections, then the outcomes of disruptive students' classmates that do not share a component with them should differ from those of the students in classrooms that lack disruptive peers. To capture this comparison, we extend $\mathbf{f}(\mathbf{\Delta}_r)\mathbf{a}_r$ to include a wedge between the classrooms that lack disruptive peers and the classrooms who have at least one of them (i.e., $\iota_r \mathbf{a}_r > 0$).

Furthermore, given that we will need to instrument $\mathbf{f}(\mathbf{\Delta}_r) = \mathbf{f_1}(\mathbf{1}[\mathbf{\Delta}_r \leq \kappa]) + \mathbf{f_2}(\mathbf{1}[\mathbf{\Delta}_r > \kappa])\mathbf{a}_r$ separately, we favor parsimonious choices of those functions. We recognize that parsimonious choices of $\mathbf{f}(\mathbf{\Delta}_r)\mathbf{a}_r$ may render the approximation to $\beta_{\bar{x}}/\beta_{\bar{y}}\mathbf{b}(\mathbf{\Delta}_r, \mathbf{G}_r)$ less precise especially for large k. However, empirical analyses of social networks have established that both the average shortest distance between pairs of agents and the diameter (the maximum distance between any pair of agents) of a social network are usually small (Jackson and Rogers, 2007). Thus, we assume linearity in $\mathbf{f_1}$ and $\mathbf{f_2}$. This linearity assumption—stemming from the need of parsimonious modelling choices due to a finite number of instruments—*does not imply* a decaying relation between $\mathbf{a}_r$ and $\mathbf{y}_r$. The parameters describing that relation are *free to be non-negative.* Our empirical model of the diffusion of disruptive peer effects becomes:

$$\mathbf{y}_r = \beta_0 \iota_r + \beta_{\mathbf{1}[.]} \mathbf{1}[\iota_r \mathbf{a}_r > 0] + \gamma_1 \tilde{\mathbf{\Delta}}_r \mathbf{a}_r + \gamma_2 \iota_r \mathbf{1}[\mathbf{\Delta}_r \mathbf{a}_r > \kappa] + \varepsilon_r, \tag{8}$$

where $\tilde{\delta}_{ij} = \delta_{ij}$ if $\delta_{ij} < \kappa$ (i.e., if i and j share the network component) and zero otherwise. $\beta_{\mathbf{1}[.]}$ captures the average effect of having at least one disruptive peer in the classroom over not having any, γ_2 captures any additional effect not connected disruptive classmates (i.e., those outside one's network component) might have, and γ_1 captures the rate at which the effect of connected disruptive peers changes depending on how far away in the social network they are located.

3.2 Identification

There are four possible threats to the identification of causal estimates of equation (8). First, common to most studies on peer effects, the existence of a feedback loop in which just as student i affects student j, j will also affect i. Second, the potential non-random sorting of disruptive peers into classrooms within schools. Third, that peer groups themselves may be endogenous, resulting from students' subjective choices of friends. Fourth, possible mechanical links between individual outcomes and peer group characteristics. Failing to address these concerns will deem our estimates mere correlations.[10] In what follows, we present an identification strategy that deals with the four threats to the identification.

[10] In particular, the third and fourth outlined identification concerns relate to the problem that Manski (1993) calls 'correlated effects' as correlations in cognitive outcomes may indeed reflect prior similarities between individuals rather than any distuptiveness' direct peer effect (contextual effect according to Manski's taxonomy) or its indirect effect through its influence on peers' cognitive outcomes (endogenous effect according to (Manski, 1993)).

3.2.1 Exogeneity of disruptiveness

We first deal with the simultaneous determination of outcomes within a peer group, which arises as a problem in our context because disruptiveness could very well be endogenous to the classroom composition. For instance, Sarzosa (2021) finds that the incidence of bullying is higher in classrooms where there are more children with uncommon traits relative to those of their classmates. Therefore, it is not possible to disentangle the effect a disruptive child has on its classmates from the effect the classmates have on the child that becomes disruptive. We address this simultaneity issue by proxying peers' disruptiveness with an exogenous measure that the student herself cannot influence: having parents who were investigated for abuse and neglect. In this sense, we follow Carrell and Hoekstra (2010), who proxy disruptive peers as children who at some point experienced domestic violence. In the same way as Carrell and Hoekstra (2010) we interpret the negative spillover in the broad sense of how some "bad apples" in the classroom may harm the academic performance of other students. We argue that our measure is exogenous because a feedback loop is very unlikely. For it to exist, a child's classmate should be able to induce abuse and neglect within the child's own household. We test the existence of a feedback loop by regressing the student's own exposure to abuse and neglect against share of male and share of female classmates that have experienced abuse and neglect while controlling for school fixed effects (Guryan et al., 2009). As expected, Table A.2 shows that there is no significant association between the two variables.

3.2.2 Random sorting into classrooms

Our second threat to identification considers the possibility of sorting into classrooms. Here, we take advantage of the institutional framework that administered student assignments to Swedish primary schools (grades 1 to 6) in the 1960s. In particular, students attended the nearest school in the neighborhood, and tracking based on ability or background was not allowed in the first six grades of the newly reformed comprehensive school based on the accord of the 1957 School Committee (Husen, 1961; Paulston, 1966; SOU1961:30, 1961).[11] Further, a homogenous curriculum and a fixed number of weekly hours of instruction (34 hours in grade 4 and 35 hours in grades 5 and 6) resulted in what explicitly was called a primary school system absent of any "organizational differentiation" of students with respect to ability or social background (SOU1961:30, 1961). The only homogenous groupings allowed during the first six grades of Swedish comprehensive school were special education classes for students with special needs (Husen, 1961). We empirically show that this was the case—at least with respect to the share of abused and neglected (A&N) students. Tables A.3 and A.4 document that the share of A&N peers in the classroom and that having at least one A&N student in the classroom are both uncorrelated with several important background characteristics (e.g., gender, being social aid recipient, birth weight, mother's age at birth, parent's ownership status of the dwelling, size of the dwelling, having older siblings) once we control for school fixed effects.

Our strategy deals with potential sorting of students into classrooms by exploiting within-school variation in classroom composition as in the strategy originally proposed by Hoxby (2000). The identifying assumption is that all other determinants of long-run outcomes are orthogonal to this within-school across-classrooms variation in peers' family circumstances and peers' parents' behavior. The intuition being that while a student by chance ends up in a classroom with, say 3 percent students with abusive parents, the student ending up in a classroom next door in the same corridor may

[11]In 1960s Sweden, students attended a single-tracked comprehensive school with six years of primary school and three years of upper secondary school. Primary education comprised two stages: grades 1 to 3 of lower primary school (from age 7 to 10) and upper primary school up until completion of sixth grade (at age 13). Formal grading of students only began in seventh grade—this substantially mitigates concerns of any informal tracking of students based on ability. The government centralized school funding, and the children's residence determined school assignments. Swedish school districts differ very much from U.S. school districts; their principal task is to allocate teachers across classes within the district. The catchment areas of the school district were determined by the maximum traveling distance to the district's lower secondary school. Recommendations of maximum traveling distance were more restrictive for primary school students, and hence there were typically more primary schools than secondary schools in one district (Fredriksson et al., 2012).

be exposed to 6 percent of students with abusive parents. We empirically confirm that parental abuse and neglect does not drive students' assignment to classrooms by contrasting our observed within-classroom proportion of A&N students with ones that would result from a random process as in Lavy et al. (2011). In a nutshell their procedure as adapted to our context goes as follows. We randomly generate the A&N status of the students in each school using a binomial distribution with a probability of success equal to the proportion of A&N students in the school. Based on the simulated data, we then calculate the within-school standard deviation of the proportion of A&N students. We repeat this process 1,000 times. These simulations allow us to compute empirical confidence intervals for the standard deviation for each school. Our Monte Carlo simulations indicate that the observed within-school variation in the proportion of A&N peers is consistent with a random process—despite having much less data than Lavy et al. (2011) to implement the procedure, as we only observe one cohort. We find that 96 percent of schools had a standard deviation of the proportion of A&N students that fell within the 95 percent confidence interval. Furthermore, Figure A.1 in the appendix shows that the distribution of the within-school standard deviation of the proportion of A&N students that we observe in our data closely follows that produced by a random process.

3.2.3 Dealing with endogneous friendship formation

The third threat to identification relates to the fact that people do not befriend randomly. In fact, socially generated networks share empirical regularities that are incompatible with random formation of links (Jackson and Rogers, 2007). Thus, friendship formation and social networks are endogenous (Carrell et al., 2013). Friendship is an active choice and, in consequence, the shape of a given network is going to be the result of those choices. This confounding problem is the sub-classroom analogue to the *correlated effects* identification issue (Lin, 2010). We address this issue by modeling the friendship formation process using a dyadic regression that is standard to the network formation literature (see the handbook chapters of Graham (2020) and de Paula (2020) for various examples).

We assume dense networks, in the sense that the probability of any two students within one forming a link is bounded away from zero, and model friendship link formation based on homophily and unobserved degree heterogeneity following Graham (2017). People with homophilous preferences tend to befriend others with similar characteristics. Homophily is a common feature of human social networks (McPherson et al., 2001; Jackson, 2010; Attanasio et al., 2012; Graham, 2017; Boucher et al., 2022). Formally, let $Z_{ij} = (\sum_{\rho=1}^{R}(Z_i^\rho - Z_j^\rho)^2)^{\frac{1}{2}}$ be the distance between i and j in K dimensions of characteristics. Student i befriends j if there is positive value of doing so: $d_{ij} = \mathbf{1}(Z'_{ij}\beta_Z + U_{ij} > 0)$, where U_{ij} is an unobserved component affecting link surplus. Then, homophily implies that dyads in which Z_{ij} is low have higher surplus implying a negative sign of β_Z.[12]

Unobserved degree heterogeneity captures another common feature of human social networks, namely that some people make friends with greater ease than others. Using the surplus structure proposed by Graham (2017), we write the unobserved component contributing to the link surplus, U_{ij}, as a function of unobserved student-level degree heterogeneity θ_i and θ_j and an idiosyncratic component ξ_{ij} drawn from a random distribution with full support, $U_{ij} = \theta_i + \theta_j + \xi_{ij}$. Moreover, for the networks to be dense, we need to assume that the function $Z'_{ij}\beta_Z + \theta_i + \theta_j$ is bounded away from zero (Johnsson and Moon, 2021). We further assume that $\eta_i = \theta_i + \nu_i$ as compounded error terms where $Z_{ij} \perp (\theta_i, \theta_j)$ and $\xi_{ij} \perp \nu_i$. Then, assuming that ξ_{ij} follows a logistic distribution and that they are independently and identically distributed across dyads,

[12] By focusing on homophily, we shut down the other channel through which friends form clusters, namely strategic friendship formation—that the utility an individual attaches to a particular friendship link depends on the presence (or absence) of other links in the network (Jackson and Wolinsky, 1996; Graham, 2015). This simplifying assumption gains tractability in two ways. First, a model that incorporates strategic aspects while allowing for agent heterogeneity would require panel data on networks, which we do not have access to (Graham, 2017). Second, opening up for strategic aspects of link formation would complicate the analysis considerably due to the possibility of multiple equilibria of the network formation model (Sheng, 2020; Badev, 2021). Graham (2015) notes that a fixed effects model of degree heterogeneity is likely to predict the link probability well even if the true link formation process happened to include strategic aspects.

we can write the likelihood of observing network $\mathbf{d}$ as

$$\Pr(\mathbf{D}=\mathbf{d}|\mathbf{Z},\theta)=\prod_{i\neq j}\left[\frac{1}{1+\exp(Z'_{ij}\beta_Z+\theta_i+\theta_j)}\right]^{1-d_{ij}}\left[\frac{\exp(Z'_{ij}\beta_Z+\theta_i+\theta_j)}{1+\exp(Z'_{ij}\beta_Z+\theta_i+\theta_j)}\right]^{d_{ij}} \tag{9}$$

Equation (9) implies that conditional on homophily $\mathbf{Z}$ and degree heterogeneity θ, links form independently.

Note that although we allow for the individual-level characteristics Z_i underlying our dyadic-level homophily variables Z_{ij} to be correlated with the error term of the individual-level outcome equation (8), the dyadic-level variables, Z_{ij}, are assumed to be uncorrelated with it, $Z_i \not\perp \eta_i$. This is because individual-level characteristics affect links through absolute values of differences $|Z_i - Z_j|$, while they affect outcomes directly. In this sense, equation (9) provides us with instruments for the endogenous variables in individual-level outcome equation (8) that are motivated through functional form. The rationale for our identifying assumption is as follows. Take for instance the distance between student i and j in health endowments at birth. If this distance is small (i.e., health endowments are similar), the students are more likely to be friends ($d_{ij}=1$) than if this distance was large (i.e., if the students had little similarity in health endowments). But conditional on school fixed effects, the distance between i and j's health endowments should not *directly* affect outcomes of either i or j (i.e., the fitted link probabilities satisfy the exclusion restriction). As such, the dyad-level variables are excluded from the outcome equation that models action choices: the difference in dimensions of the dyad-level link formation equation (9) and the individual-level outcome equation (2) provides an exclusion restriction that relies on nonlinearities (Patacchini et al., 2017; Hsieh and Lee, 2015). Sub-section 3.3 describes the exact instrumental variable estimation strategy.

3.2.4 The Angrist critique

In addition to the identification considerations described above, we must take into account Angrist (2014)'s critique that flags for the *mechanical* correlation that arises between own and peer characteristics in estimates relating individual outcomes to peer group averages. Angrist (2014) shows that excluding individual i from the peer group average on the right hand side (RHS) of a grouping estimator only partly solves the mechanical link between both sides of the estimating equation, and ever the less so the larger the group size. He also shows that random group formation will not help solving this spurious nature of peer effects estimated using a leave-out grouping estimator.[13] To ameliorate this issue, we follow Angrist's recommendation to separate those who are potential subjects to the peer effects from the senders of the effect. We do so by dropping the abused and neglected students. That way, we also exclude the analysis of the effect of abuse and neglect by own parents which is not the focus of this paper.

3.3 Estimation

Link formation. To estimate equation (9), we arrange the data in dyadic form considering the possible friendships with every other student in the classroom. We estimate equation (9) with a logistic regression that includes fixed-effects of the students sending and receiving the friendship nomination intended to capture the unobserved degree heterogeneity (we model undirected social networks), where the dyadic covariate space Z_{ij} contains variables indicating whether i and j have the *same* gender, have parents with the same social aid recipient status, parents with dwelling ownership, and father's

[13] Angrist (2014) shows first algebraically that the estimated coefficient of a grouping estimator differs by a factor of $\frac{N}{N-1}$ according to whether or not student i is excluded from the RHS peer group average, meaning that the difference between the two peer group averages will become small once N is large. Angrist (2014) goes on to show in the 2SLS framework (allowing the first stage regression to do the averaging of the RHS), that if groups are formed randomly (i.e., if the first stage is weak), a bivariate grouping estimator that leaves-out student i from the group average would mechanically produce strong negative estimates even in the absence of any peer effects.

occupational state at the time of birth, if they lived in the same neighborhood block in 1953, the *distance* between i and j with respect to two neonatal health indices. We report the estimation results in Table I.2 of Online Appendix I. The results show statistically significant homophily with respect to most aforementioned variables.

Diffusion model. In our main estimating equation (8), we instrument i's distance to an A&N peer in i's network component, $\tilde{\boldsymbol{\Delta}}_r\mathbf{a}_r$, and the indicator function that captures possible effects of A&N peers who are outside of i's component, $\mathbf{1}[\boldsymbol{\Delta_r}\mathbf{a}_r > \kappa]$, with a set of predicted propensities of friendship based on homophily $\mathbf{Z}$ that come from our estimations of (9). Specifically, we use two instruments: i) the probability that student i forms a link with the A&N peer q, $\hat{\Pr}(D_{iq} = 1|Z_{iq}, \theta_i, \theta_q)$, and ii) the probability that student i's friends are friends with the A&N peer q while i and q are not friends.[14] This instrumental variable solution forms a system of three equations that we estimate using Limited Information Maximum Likelihood estimation (LIML).

As an alternative to the LIML estimation, we propose a more parsimonious *empirical* operationalization of equation (7) in which we top-code the distance to the A&N peer of those students that are not socially connected to her. That is, we impute the distance to the A&N peer of the disconnected students with the maximum observed distance of those connected to the troubled peer plus one. Hence,

$$\boldsymbol{\Delta}'_r = \begin{cases} \boldsymbol{\Delta}_r\mathbf{a}_r & \text{if } \boldsymbol{\Delta}\mathbf{a}_r \leq \kappa \\ (\kappa+1)\mathbf{a}_r & \text{if } \boldsymbol{\Delta}\mathbf{a}_r > \kappa \end{cases}$$

That way, the estimating equation only contains one endogenous variable, namely $\boldsymbol{\Delta}'_r$.

$$\mathbf{y}_r = \beta_0\iota_r + \beta_{\mathbf{1}[.]}\mathbf{1}[\iota_r\mathbf{a}_r \geq 0] + \gamma_1\boldsymbol{\Delta}'_r + \varepsilon_r,$$

[14]See Kelejian and Piras (2014) and König et al. (2019) for applications that fit the complete adjacency matrix by exploiting the same difference in dimensionality between the dyadic-level link formation model and the outcome equation.

A caveat to this top-coding approach is that it might place some of the disconnected classmates closer to the A&N peer than they would actually be if we observed an untruncated classroom network. For this reason, we consider the LIML model of (8) our main estimating equation. We further deal with additional consequences of network truncation in Section 6.1.

4 Main results

This section presents the main results of the model outlined in subsection 3.1.2 on the diffusion of the negative peer effect with social distance in terms of shortest path length within the network to the A&N peer.[15] Table 2 reports the results of estimating equation (8) using the limited information maximum likelihood (LIML) estimator and the estimator that top-codes (TC) the distance to not-connected peers to the network diameter plus one. The results show that the network architecture is an essential determinant of the size of the negative externality exerted by troubled classroom peers. Both models tell a consistent story: score losses caused by having an A&N peer diminish as social distance to her increases. The top panel shows that each edge away from the troubled peer reduces the score losses by 0.49 points which corresponds to 8 percent of a SD in verbal scores, 0.53 points (7 percent of a SD) in numeric scores, 0.47 points in spatial scores (7 percent of a SD) and 8 percent of a SD in marks at ninth grade. Thus, the direct friends (i.e., those with a social distance equal to one edge from the

[15] In online appendix II, we document the presence of overall negative spillovers of disruptive peers. We estimate the average total effect using the reduced form of the typical LIMM as in Carrell and Hoekstra (2010) and Carrell et al. (2018). Table II.1 shows that peer abuse and neglect has a significant and substantial effect on own verbal and numeric abilities in sixth grade and on GPA in ninth grade. In general, the inclusion of one additional A&N classmate to a classroom of 20 students decreases verbal and numeric ability in sixth grades by 3.2 percent and 3.0 percent of a SD respectively. Our results are very much in line with the results of Carrell and Hoekstra (2010) who explore whether children exposed to domestic violence exert negative externalities on their classmates' math and reading test scores for third to fifth graders in a county in Florida. In Table II.1, we also examine the results by gender of the peer exposed to abuse and neglect. We find that while A&N males affect their male and female classmates, A&N females have negative externalities on other females but not on males.

Table 2: Effects of Distance to Abused and Neglected Peers on Cognitive Scores at Ages 13 and 16. Full Sample

	IQ Components at 13						*Grades at 16*	
	Verbal		Numeric		Spatial		GPA in grade 9	
	TC	LIML	TC	LIML	TC	LIML	TC	LIML
A&N in class	-1.166	-1.517	0.916	-1.446	-0.688	-1.105	-0.136	-0.202
	(0.462)	(0.619)	(0.586)	(0.649)	(0.524)	(0.779)	(0.078)	(0.108)
Not connected		1.442		1.347		1.319		0.185
		(0.566)		(0.677)		(0.781)		(0.107)
Dist. to A&N	0.259	0.493	0.223	0.529	0.221	0.473	0.036	0.079
	(0.122)	(0.218)	(0.154)	(0.263)	(0.138)	(0.294)	(0.020)	(0.042)
Observations	7,464	7,464	7,464	7,464	7,464	7,464	7,101	7,101
Effect by Distance:								
One edge	-0.907	-1.024	-0.693	-0.917	-0.467	-0.631	-0.100	-0.123
	(0.353)	(0.425)	(0.448)	(0.427)	(0.400)	(0.520)	(0.060)	(0.070)
Two edges	-0.648	-0.531	-0.470	-0.388	-0.246	-0.158	-0.064	-0.044
	(0.256)	(0.269)	(0.325)	(0.284)	(0.290)	(0.325)	(0.043)	(0.042)
Three edges	-0.389	-0.039	-0.247	0.141	-0.025	0.315	-0.028	0.035
	(0.190)	(0.242)	(0.241)	(0.342)	(0.215)	(0.337)	(0.032)	(0.046)

Note: Data from Stockholm Birth Cohort Study. Sample excludes classes with less than seven students, schools with only one class in grade 6, schools with special education classrooms and classrooms in which no one participated in the sociometric survey. Sample restricted to individuals who were not abused and neglected (Angrist, 2014). *A&N in class* takes on value one if there is a student in the classroom whose parents underwent an investigation for abuse and neglect by the child protection services (CWC) that was substantiated and zero otherwise; *Not connected* takes on value one if the potential disruptive peer in the classroom does not belong to the social network of the student; and *Dist. to A&N* stands for the path length between the student and the closest disruptive peer. All regressions include school fixed effects, a dummy for being female, whether family receive social assistance, birth weight, mother's age, dwelling type, dwelling ownership and number of older siblings. Further, the estimated degree heterogeneity $\hat{\theta}_i$ for student i is treated as an included instrument and is hence included in both the first stages and the second stage. Columns labeled TC report the estimates of model in which the distance between two not-connected peers is top-coded to be equal to the network diameter plus one. Columns labeled LIML report the estimates of the distance model outlined in equation (8). The bottom panel reports the linear combination of the estimates associated to whether there is an A&N peer in class and the distance to the closest A&N peer, when evaluating the latter values from one to three. Standard errors, shown in parentheses, are clustered at the school level.

disruptive peer) suffer the greatest losses in terms of cognitive ability. The effect then dissipates following an inverse distance rule.

The bottom panel of Table 2 is intended to aid interpretation by reporting the estimated effects at different values of the geodesic to the A&N peer. It shows that verbal and numeric ability test scores decrease on average by roughly -1.02 ($\approx$ -1.52+0.49) points and -0.92 ($\approx$ -1.45+0.53) points respectively among direct friends of A&N classroom peers. They imply a reduction in verbal and numeric test scores by 16.8 percent of a SD and 12 percent of a SD, respectively. We find a similar pattern when analyzing the effects of having an A&N peer on student performance (Spring term GPA) three years later in 9^{th} grade. The closest friends of the disruptive peers will see their grades fall by 12.3 percent of a SD. For the classmates not directly connected to the A&N peer, the negative effect on cognitive outcomes wanes as the distance increases. For instance, the friends of A&N peer's friends (friends of friends) suffer a decrease of -0.53 ($\approx$ -1.52+2*0.49) points in the verbal score and -0.39 points in the numeric score. Our results show that those students who are three edges away from the A&N peer incur almost no harm of being associated with her socially.

Although the effect of A&N peers on their classmates' spatial ability also follows the inverse distance pattern, it is relatively small even for the closest friends and wanes rapidly with distance. The A&N peers' closest friends lose roughly 0.63 points on their spatial scores—but the effect is statistically indistinguishable from zero. The reason behind the effect being weaker for spatial ability may be that the spatial ability component of the intelligence test is weighted more towards fluid intelligence. Fluid intelligence is considered less responsive to external inputs, as discussed in Section 2.

Our estimates of the effects of A&N peers on those who are socially disconnected from them provide further evidence of the importance of the social network's architecture in diffusing the externality. Our results show that regardless of the outcome we measure, not being socially connected to the A&N peer, as measured by undirected links, offsets almost entirely the negative effect of having her in the classroom. In other words, unlike those socially close, A&N classmates have no adverse effect on peers who belong to a different network component. Taken together, our results indicate that A&N

peers' disruptiveness does not affect everyone equally in the sense of affecting overall learning through interrupting during class or capturing resources from the teacher. If that was the case, even disconnected peers would be affected. Instead, the effects of disruptiveness concentrate on the inner social circle of the disruptive students.

Table 3 presents the subgroup analysis by gender. It shows that social closeness to an A&N peer is more harmful to males' language development than to that of females. Each additional edge away from the A&N peer reduces its impact on males' verbal scores by 0.43 points or 7 percent of a SD. On average, A&N peers' close male friends suffer a 1.3 points decrease in their verbal score. The male friends of those friends lose 0.88 points. The evidence of a social distance-induced fade out of a disruptive peer effect on verbal ability is substantially weaker and statistically insignificant for females. As to numeric ability, we find a more similar pattern of effect heterogeneity by social distance for females and males; each additional edge away from the A&N peer reduces the harm conferred on males' numeric scores by 0.49 and on females' numeric scores by 0.63 points. The two parameters are not statistically different form each other. The bottom panel of Table 3 indicates that according to the LIML estimation, males lose one point due to their A&N close friend, and their male friend who are two edges away from the A&N peer (and thus form an intransitive triad with the former two), lose in turn 0.55 points. Female friends of A&N students lose 0.71 points and those two edges away from the A&N peer do not suffer a noticeable effect.

The results on spatial ability also show differences by gender. Although not statistically significant at standard levels, our results suggest that the closest male friends of the A&N peer lose about 0.87 points in the spatial score. Interestingly, we find a significant negative effect that persists until 9^{th} grade only for females. We find the effect also decays with social distance. Close female friends of the A&N peer lose around 22.3 percent of a SD in their 9^{th} grade marks. Female friends of friends of the A&N peer lose about 8 percent of a SD.

Taken together, these results suggest that A&N peers affect females and males differently depending on the cognitive dimension explored. On the one hand, the sheer existence of A&N peers affects females' verbal test scores regardless of their gender or relative position in the social network. On the other hand, only male A&N peers affect males *and only* if they are in their inner social circle. This difference is remarkable especially considering that the set of A&N peers comprise females and males in equal proportions—being abused and neglected depends on the parents' and not the children's characteristics. Therefore, females have A&N peers in their inner circle as much as males do. Five percent of males and five percent of females have at least one A&N peer among their closest friends. Thus, despite both genders being equally well connected with A&N peers, we observe different mechanisms affecting males' and females' verbal scores.

Table 3: Effects of Distance to Abused and Neglected Peers on Cognitive Scores at Ages 13 and 16, By Gender

	IQ Components at 13						*Grades at 16*	
	Verbal		Numeric		Spatial		GPA in grade 9	
	TC	LIML	TC	LIML	TC	LIML	TC	LIML
Males								
A&N in class	-1.731	-2.129	-1.131	-1.515	-1.189	-1.567	-0.105	-0.119
	(0.661)	(0.878)	(0.881)	(1.011)	(0.795)	(1.203)	(0.117)	(0.132)
Not connected		2.381		1.617		2.558		0.202
		(0.831)		(1.058)		(1.191)		(0.146)
Dist. to A&N	0.428	0.687	0.275	0.491	0.452	0.701	0.040	0.054
	(0.184)	(0.323)	(0.245)	(0.396)	(0.221)	(0.440)	(0.033)	(0.053)
Observations	3,647	3,647	3,647	3,647	3,647	3,647	3,474	3,474
Females								
A&N in class	-0.463	-0.708	-0.318	-1.395	-0.357	-1.044	-0.173	-0.368
	(0.782)	(0.818)	(0.946)	(1.013)	(0.835)	(0.938)	(0.124)	(0.158)
Not connected		0.506		1.388		0.756		0.278
		(0.832)		(1.063)		(0.987)		(0.154)
Dist. to A&N	0.074	0.214	0.111	0.628	0.088	0.456	0.037	0.145
	(0.204)	(0.311)	(0.247)	(0.406)	(0.218)	(0.363)	(0.032)	(0.061)
Observations	3,817	3,817	3,817	3,817	3,817	3,817	3,627	3,627
Males:								
Effect by Distance:								
One edge	-1.304	-1.442	-0.856	-1.023	-0.737	-0.866	-0.064	-0.065
	(0.496)	(0.587)	(0.660)	(0.665)	(0.596)	(0.802)	(0.088)	(0.086)
Two edges	-0.876	-0.754	-0.581	-0.532	-0.285	-0.165	-0.024	-0.011
	(0.349)	(0.356)	(0.465)	(0.418)	(0.419)	(0.477)	(0.062)	(0.055)
Three edges	-0.448	-0.067	-0.307	-0.041	0.167	0.536	0.016	0.044
	(0.257)	(0.342)	(0.342)	(0.471)	(0.308)	(0.446)	(0.045)	(0.066)
Females:								
Effect at Distance:								
One edge	-0.390	-0.494	-0.207	-0.766	-0.269	-0.588	-0.135	-0.223
	(0.594)	(0.539)	(0.719)	(0.648)	(0.634)	(0.608)	(0.095)	(0.103)
Two edges	-0.316	-0.280	-0.097	-0.138	-0.181	-0.132	-0.098	-0.078
	(0.421)	(0.325)	(0.510)	(0.377)	(0.449)	(0.351)	(0.067)	(0.063)
Three edges	-0.242	-0.067	0.014	0.490	-0.093	0.325	-0.061	0.067
	(0.292)	(0.336)	(0.353)	(0.441)	(0.311)	(0.375)	(0.047)	(0.068)

Note: Data from Stockholm Birth Cohort Study. See Table notes of Table 2 for sample restrictions, variable definitions and control variables. Standard errors, shown in parentheses, are clustered at the school level.

5 Understanding the nature of the peer effect

We have thus far explored the magnitude and the reach of the peer effect and will in this section turn our attention towards the nature of the peer effect. In the case of disruptiveness, it is particularly important to be informed about the extent to which it is the A&N peer's disruptiveness that affects student's cognitive achievement and student performance and to which the A&N peer might affect her classmates through her own academic achievement and student performance. In particular, we explore how the A&N peer effect on cognitive outcomes finds its path through the classroom's social network. For this purpose, we return to the network version of LIMM in equation (2) and consider it the structural formulation of the peer-effect process. The LIMM's structure allows us to inquire the extent to which the A&N peer affects classmates' cognitive achievement via her disruptiveness itself ($\beta_{\bar{x}}$, the direct contextual effect) or indirectly via the disruptive peers' lower cognitive outcomes ($\beta_{\bar{y}}$, the indirect endogenous effect).

Identification of the structural parameters in equation (2) requires addressing two issues. First, the endogeneity of $\mathbf{G}_r\mathbf{y}_r$ (i.e., the reflection problem (Manski, 1993)). Second, just like in Section 3, the endogeneity of friendship formation (i.e., the endogeneity of the row-normalized adjacency matrix $\mathbf{G}_r$). As we will explain in detail in Section 5.1.1, we deal with both identification issues by exploiting the social network architecture in each classroom. In a nutshell, we follow the joint regression framework of Johnsson and Moon (2021) and extend the 2SLS procedure that draws instruments for $\mathbf{G}_r\mathbf{y}_r$ from partially overlapping networks (Bramoullé et al., 2009; De Giorgi et al., 2010) by including a control function—that relies on the unobserved social ability (θ_i)—to account for endogenous friendship formation.

5.1 Estimating the structural LIMM

5.1.1 2SLS approach to the reflection problem

As in Bramoullé et al. (2009), we use the series expansion $(\mathbf{I}_r - \beta_{\bar{y}}\mathbf{G}_r)^{-1} = \sum_{k=0}^{\infty} \beta_{\bar{y}}^k \mathbf{G}_r^k$ to express the reduced form as:

$$\mathbf{y}_r = \frac{\beta_0}{1-\beta_{\bar{y}}}\iota_r + \beta \mathbf{a}_r + (\beta\beta_{\bar{y}} + \beta_{\bar{x}}) \sum_{k=0}^{\infty} \beta_{\bar{y}}^k \mathbf{G}_r^{k+1}\mathbf{a}_r + \sum_{k=0}^{\infty} \beta_{\bar{y}}^k \mathbf{G}_r^k \eta_r \tag{10}$$

The series expansion formulation of the reduced form LIMM in equation (10) implies that, given linear independence of I_r, $\mathbf{G}_r$ and $\mathbf{G}_r^2$ and $\beta\beta_{\bar{y}} + \beta_{\bar{x}} \neq 0$, $[\mathbf{G}_r^2\mathbf{a}_r,\ \mathbf{G}_r^3\mathbf{a}_r,\ \ldots]$ are the best instruments for $\mathbf{G}_r\mathbf{y}_r$ in equation (2).[16] The important insight made by Bramoullé et al. (2009) is that equation (10) delivers a first stage for a two-stage least squares (2SLS) approach to estimating the structural LIMM outlined in equation (2) with excluded instruments for the endogenous variable $\mathbf{G}_r\mathbf{y}_r$ based on the identifying power of intransitive triads (i.e., matrices $\mathbf{I}$, $\mathbf{G}$ and $\mathbf{G}^2$ are linearly independent). Apart from intransitivity of friendships, another key to this approach is the exclusion restriction that friends-of-friends' characteristics (in our case, their abuse and neglect) do not affect own cognitive achievement directly (Bramoullé et al., 2009; De Giorgi et al., 2010). Then, the 2SLS estimator becomes

$$\hat{\beta}_{2SLS} = \left(\mathbf{W}_r'\mathbf{C}_r\left(\mathbf{C}_r'\mathbf{C}_r\right)\mathbf{C}_r\mathbf{W}_r\right)^{-1}\mathbf{W}_r'\mathbf{C}_r\left(\mathbf{C}_r'\mathbf{C}_r\right)^{-1}\mathbf{C}_r'\mathbf{y}_r \tag{11}$$

where $\mathbf{W}_r = [\mathbf{G}_r\mathbf{y}_r,\ \mathbf{a}_r,\ \mathbf{G}_r\mathbf{a}_r]$ and $\mathbf{C}_r = [\mathbf{a}_r,\ \mathbf{G}_r\mathbf{a}_r,\ \mathbf{G}_r^2\mathbf{a}_r,\ \mathbf{G}_r^3\mathbf{a}_r]$

[16] To keep exposition simple, in this subsection we keep the school-level fixed-effects implicit. We do include them in the estimation of the structural model in order to leverage on the random allocations of students to classrooms within schools.

5.1.2 Control function solution to homophily bias/endogenous friendship formation

However, when friendship formation is endogenous, matrix $\mathbf{C}_r$ is no longer orthogonal to the error term η_r because unobserved student characteristics are potentially correlated with both friendship formation and individual outcomes. In that context, Johnsson and Moon (2021) show that one can use a control function approach based on the unobserved degree heterogeneity θ_i that is identified in the friendship formation equation (9). They treat θ_i as an individual fixed-effect that is correlated with $\mathbf{a}_r$ (e.g., disruptive children might face difficulty making friends), which they assume accounts for all the dependence between $\mathbf{a}_r$ and η_r. Based on this assumption, Johnsson and Moon (2021) show that $\mathbf{c}_i - \mathbb{E}[\mathbf{c}_i|\theta_i]$ is orthogonal to $\eta_i - \mathbb{E}[\eta_i|\theta_i]$. Then, the θ_i-partialled-out instrument matrix $\widetilde{\mathbf{C}}_r = \mathbf{C}_r - \mathbb{E}[\mathbf{C}_r|\theta]$ is a good IV for the θ_i-partialled-out regressors matrix $\widetilde{\mathbf{W}}_r = \mathbf{W}_r - \mathbb{E}[\mathbf{W}_r|\theta]$.

We operationalize this in a two-step procedure. First, we create a flexible function $\mathbf{h}(\theta_i)$ out of our estimates of unobserved degree heterogeneity for each student.[17] Following Johnsson and Moon (2021), we approximate $\mathbf{h}(\theta_i)$ with fifth- and sixth-degree polynomial sieves. From here, we go on and create $\widehat{\widetilde{\mathbf{C}}}_r$ and $\widehat{\widetilde{\mathbf{W}}}_r$ by partialling $\hat{\mathbf{h}}(\theta_i)$ out of every term therein. Second, we perform a 2SLS procedure where we instrument $\widehat{\widetilde{\mathbf{W}}}_r$ with $\widehat{\widetilde{\mathbf{C}}}_r$. Therefore, the new 2SLS estimator becomes:

$$\hat{\beta}_{J\&M} = \left(\widehat{\widetilde{\mathbf{W}}}_r'\widehat{\widetilde{\mathbf{C}}}_r\left(\widehat{\widetilde{\mathbf{C}}}_r'\widehat{\widetilde{\mathbf{C}}}_r\right)\widehat{\widetilde{\mathbf{C}}}_r\widehat{\widetilde{\mathbf{W}}}_r\right)^{-1}\widehat{\widetilde{\mathbf{W}}}_r'\widehat{\widetilde{\mathbf{C}}}_r\left(\widehat{\widetilde{\mathbf{C}}}_r'\widehat{\widetilde{\mathbf{C}}}_r\right)^{-1}\widehat{\widetilde{\mathbf{C}}}_r'\widehat{\widetilde{\mathbf{y}}}_r \tag{12}$$

5.2 Results of the structural LIMM

In Table 4, we present the estimation results of the structural model. For each outcome, we report the results of two different models of which the first one is the baseline 2SLS of

[17] Following Johnsson and Moon (2021), we use the estimates of θ_i that we obtain from the joint fixed effect structure in (9). That is, we use the friendship formation model with undirected links.

estimating equation (11) and the second one is the control function model (2SLS+CF) of equation (12). Of particular interest are the estimates on verbal and numeric ability as these are the outcomes for which we find peer effects in our diffusion model (Section 4) and in the reduced form (Section II). The findings suggest that the mechanisms through which A&N students' disruptiveness harms peers depend on the dimension of cognitive ability. While the endogenous effect is larger than the direct contextual effect in depleting verbal ability, the opposite is true for numeric ability. This difference indicates that while A&N students' disruptiveness affects students' acquisition of verbal ability through the drop in the A&N peer's verbal performance, it depletes numeric skills directly. These findings are in line with the psychometric and pedagogical literature showing that children exposed to richer vocabulary develop greater language skills (Hart and Risley, 1995), while acquiring math skills is less reliant on peer social connections and more reliant on the intrinsic ability to perform certain cognitive processes (Purpura et al., 2017).

To further explore effect heterogeneity by gender, we construct separate adjacency matrices $\mathbf{G}^M$ and $\mathbf{G}^F$ for the classroom's male and female networks, respectively.[18] We present the results in Online Appendix III. The results in Table III.1 suggest that the endogenous effect is an important pathway of A&N peers' effect on males but not on females. Furthermore, the direct effect of A&N male disruptiveness seems to substantially deter numeric skill acquisition for both genders (the point estimates on females being substantially larger, although less precise).

Interpretation of the estimates in the structural model is not straight forward. First, the estimates are embedded in the interconnections within social clusters (i.e., friends have friends in common) that may create a feedback loop. Second, matrix $\mathbf{G}$ is row-normalized (i.e., $\mathbf{G} = \text{diag}(\mathbf{M})^{-1}\mathbf{D}$) so that the interpretation of the estimates depends

[18] In theory, splitting the adjacency matrix into gender-specific matrices might violate the exclusion restriction of the 2SLS strategy as the omission of opposite-sex friends from the sample will censor same-sex friends of friends linked to each other by an opposite-sex friend. However, as noted above, cross-gender friendships are rare in our data. Thus, the gender-censoring induced bias of the estimates is likely to be small.

Table 4: Structural Estimation Results

	IQ Components at 13						*Grades at 16*	
	Verbal		Numeric		Spatial		GPA in grade 9	
	2SLS	2SLS+CF	2SLS	2SLS+CF	2SLS	2SLS+CF	2SLS	2SLS+CF
Gy	0.640	0.635	0.417	0.426	0.628	0.656	0.312	0.332
	(0.211)	(0.283)	(0.238)	(0.327)	(0.624)	(0.825)	(0.364)	(0.373)
a	-1.266	-1.264	-1.553	-1.538	-0.866	-0.859	-29.189	-28.950
	(0.317)	(0.366)	(0.496)	(0.507)	(0.344)	(0.424)	(5.157)	(5.176)
Ga	-0.329	-0.397	-1.702	-1.736	0.179	0.184	-14.836	-14.235
	(0.875)	(0.877)	(0.895)	(1.189)	(1.113)	(1.151)	(17.519)	(16.510)
Observations	7,964	7,964	7,974	7,974	7,964	7,964	7,878	7,878

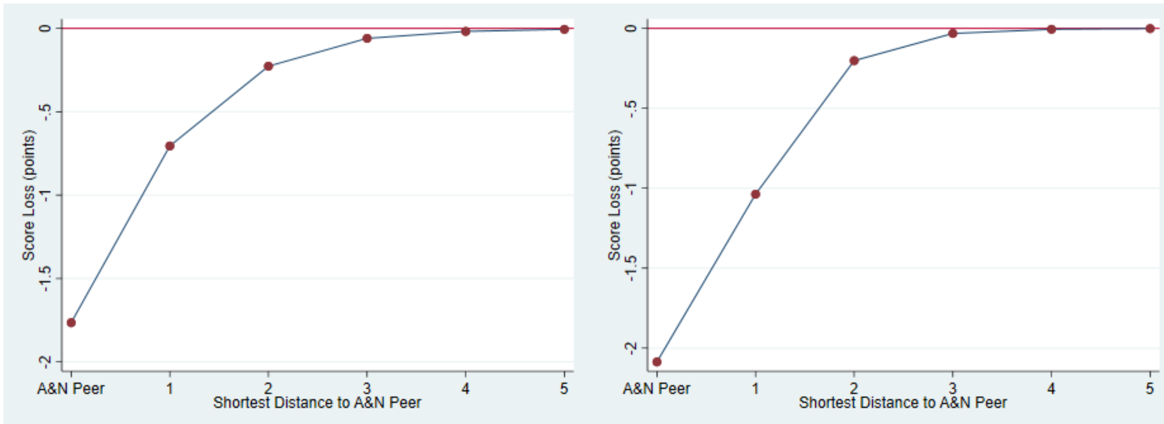
Note: Data from Stockholm Birth Cohort Study. Table present the estimates of structural model (2). The coefficients associated with **Gy** and **Ga** represent the endogenous and contextual components of the peer-effects respectively. Estimations include school fixed effects. Sample excludes classes with less than seven students, schools with only one class in grade 6, schools with special education classrooms and classrooms in which no one participated in the sociometric survey. Standard errors in parentheses. 2SLS models use clustered standard errors at the classroom level. 2SLS+CF models use clustered at the classroom level bootstrapped standard errors.

on the total number friends. For instance, our estimates indicate that a student's verbal score drops by 0.64 if all her friends lost one point in their own verbal score. Alternatively, she would lose 0.32 if only half of her friends lost one point in their own verbal score. Thus, excluding feedback loops, a drop of 1.26 points in the verbal score of a student due to her own disruptiveness results in a $0.8/m$ points drop in her immediate friends' verbal scores.

To aid interpretation of the LIMM's estimated parameters, we take a network, assign A&N peers within it and, based on the estimated parameters and the friendship links, let the model—and therefore, the location relative to the A&N peer—yield the expected outcome for each individual. That way, we let the effects work their way through the friendship loops and cliques that produce feedback. Of course, the shape of the network matters in such an exercise. For this reason, we take 1,000 random draws of social networks that, on average, match the characteristics of the average classroom in our data. Specifically, we calibrate a distribution of dyad linking probabilities yielding networks that, in expectation, have the same density, number of edges and clustering as

Figure 1: Structural Model: Results From Simulations Overall Sample

(a) Verbal Score (b) Numeric Score



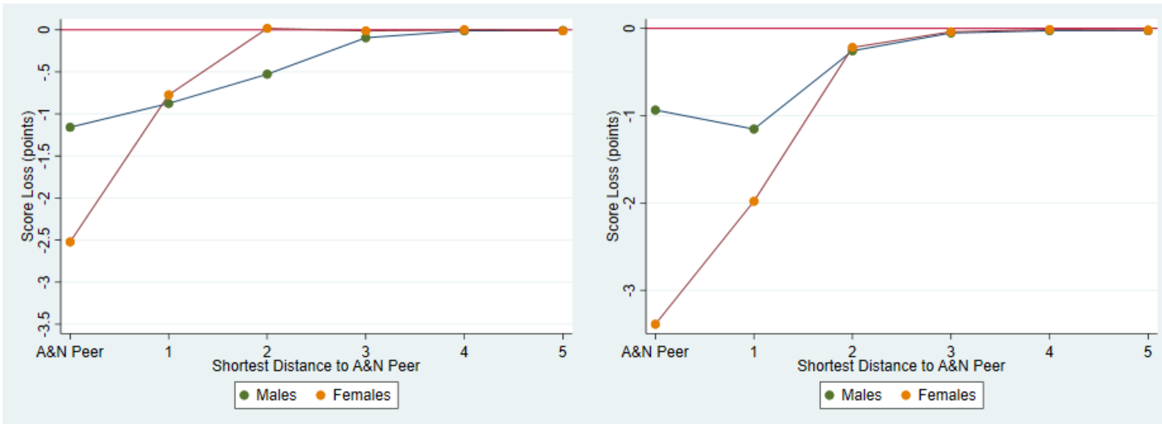
Note: Data from Stockholm Birth Cohort Study. Figures present the score losses by the distance to an A&N peer according to the LIMM as estimated using 2SLS+CF estimator reported in Table 4. We simulated 1000 networks in which the dyad linking probabilities were such that the networks matched, on average, the number of edges and the clustering of the average classroom network we observe in the data. See mode details on the simulation in Online Appendix IV.

the average classroom in our data. Within each draw, we randomly set two nodes in the network to be A&N students—thus, effectively randomizing the A&N peers' location in the network. Using the parameters reported in Table 4, we quantify the effect inflicted upon each node by the A&N peers. We then summarize the results of our simulations by aggregating the effects by the distance to the closest A&N student. We present further details of the Monte Carlo simulation in Online Appendix IV.

Figures 1 present the results of the simulations for the verbal and numeric ability dimensions. They show that even though the structural LIMM estimates suggest that the effects of A&N peers on each dimension of ability go through different channels, they dissipate at a rate proportional to the distance to the A&N peer. Confirming the results of our diffusion model in Section 4, the fade-out rates are such that there is no longer a detectable effect beyond two friendship links. The effect is larger for the A&N students themselves. They not only suffer the direct effect of exposure to abuse and neglect (i.e., **a**) but also a feedback effect that their harmed friends circle back to

Figure 2: Structural Model: Results From Simulations Gender-Specific Effects

(a) Verbal Score (b) Numeric Score



Note: Data from Stockholm Birth Cohort Study. Figures present the score losses by the distance to a disruptive peer according to the LIMM as estimated using 2SLS+CF estimator reported in Table III.1. We simulated 1000 networks in which the dyad linking probabilities were such that the networks matched, on average, the number of edges and the clustering of the average classroom network we observe in the data. See mode details on the simulation in the Appendix.

them. We find that the direct contextual effect of the A&N peers on their friends (i.e., **Ga**) is larger for numeric ability than verbal ability. However, the endogenous effect is smaller for numeric than for verbal, and thus—in accordance with our distance model results—the effect wanes faster for numeric than verbal ability as the social distance to the A&N peer increases.

Using a similar procedure, we also interpret the gender-subgroup estimates presented in Table III.1. That is, we randomly generate 1,000 networks that in expectation, have the same density, number of edges and clustering as the average classroom in our data. We randomly select two nodes to be A&N students. We assume that nodes are evenly distributed by gender and randomize both the network location and the gender of the A&N nodes. We plot the simulation results aggregated by the distance to the A&N node of the same gender in Figures 2. In congruence with the gender-specific distance results presented in Table 3, Figures 2 show that while the effects on verbal ability dissipate throughout the network differently for females and males, the treatment propagation on numeric ability seems to be more similar across genders. In particular, a male friend of the friend of a male A&N peer still loses some verbal score points, while this peer effect does not reach friends of friends of a female A&N peer. As to numeric ability, although A&N females lose more than their male counterparts, and exert a greater externality for their friends, the gender difference is leveled for the friends of the A&N peer's friends.

5.3 Counterfactual policy exercises

Taking advantage of the structural results, we consider two counterfactual scenarios that evaluate possible policy interventions. First, in the spirit of Díaz et al. (2021), we remove one A&N peer from the classrooms (we call this counterfactual "*One A&N less*"). This can be seen as a social intervention that reduces abusive behaviors in parents (e.g., reduction in the incidence of alcoholism in adults). Second, we consider a case where policy makers want to reduce the cognitive achievement losses due to the presence of A&N peers. They are faced with two alternatives: a) target resources on the A&N peers (e.g., hire tutors or organize counseling for the A&N peers), and b) spread out the same resources across all members of the classroom (i.e., use the same amount of tutoring or counseling hours but divided equally among all classmates).[19] To operationalize these alternatives, we consider a typical 22-student classroom with two A&N peers. For a), the policy maker considers a remedial intervention like the one described in Özek (2021), who finds that having under-performing middle school students take a reading remedial class in addition to their normal course load boosts their scores by 11 percent of a SD relative to similarly-skilled ineligible students (we call this counterfactual "*Remedial A&N only*"). As a comparison, in b), the policy maker

[19]We refrain from considering policies based on allocative designs (e.g., equalizing the share of disruptive peers across classes) as such prescriptions would require knowledge about the presence of social multipliers, the identification of which is not straightforward. Some research has equated the endogenous effect with a social multiplier. However, recent literature has shown that a significant non-zero coefficient of $\beta_{\bar{y}}$ need not be evidence of a social multiplier (Bramoullé et al., 2020; Boucher and Fortin, 2016).

"subsidizes" achievement of all 22 students by a 0.5 percent of a SD improvement relative to a no-intervention scenario (we call this counterfactual "*Remedial All*").[20] We think of this scenario as an intervention where all students get one tutoring session per-month as opposed to enrolling on a full course like in the "Remedial A&N only" option. Our key assumption in order to be able to compare them is that the score returns to a dollar spent are linear and equal across students.

Each line in Figure 3 reports the score losses by distance to the closest A&N peer for each counterfactual scenario. Figure 4 represents the corresponding total score losses relative to the actual loss based on our baseline results of the structural model. Each line and bar are the results of 1,000 simulations under each counterfactual scenario. We also present the actual effects estimated in our structural model for comparison. The figures show that, among the three proposed policies, the one reducing the number of A&N students in the classroom has the smallest impact in palliating the negative effect of A&N peers. At face value, it might seem counterintuitive. However, note two important features of our setting. First, children at this age almost exclusively befriend children of the same gender (Stehlé et al., 2013)—in our data, 97 percent of friendship nominations are within gender. Second, A&N students are roughly evenly split across genders, a pattern coming directly from the circumstance that parental abuse and neglect is equally prevalent among parents of daughters and sons. Therefore, dropping one of the two A&N students from the classroom implies that it is likely that students from a different gender than the removed student will remain connected to the remaining A&N student. Thus, the students closest to the remaining A&N peer are largely unaffected. The total effects' small relative reductions of 10.3 and 9.6 percent

[20] We consider the 0.5 percent shift to be an upper bound. Swedish statistics indicate that the average per-pupil cost of a school year is SEK 67,000. If we consider that 6^{th} grade students take 10 subjects in a school year (e.g., Swedish, English, math, science, social studies, etc.) and that the average classroom has two A&N peers, then the "Remedial A&N only" policy would spend SEK 609 ((67,000*2)/(10*22)) per-pupil. That represents 0.91 percent of the per-pupil yearly cost of a student. Estimates using Swedish data contemporaneous to ours indicate that the causal effect of an extra year of schooling on scores amounts to 17 percent of a SD (Meghir et al., 2013). Thus, a subsidy of SEK 609 can *buy* a boost in scores of about 0.155 percent (0.91%*17%) of a SD for everyone in the classroom. That is less than a third of the 0.5 percent of SD shift in scores that we use in our simulation exercises.

Figure 3: Structural Results: Effect of A&N in Classroom by Distance to Peer

(a) Verbal (b) Numeric

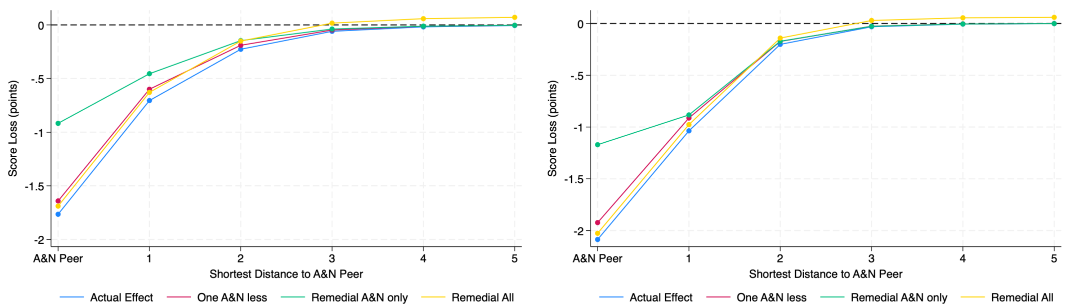
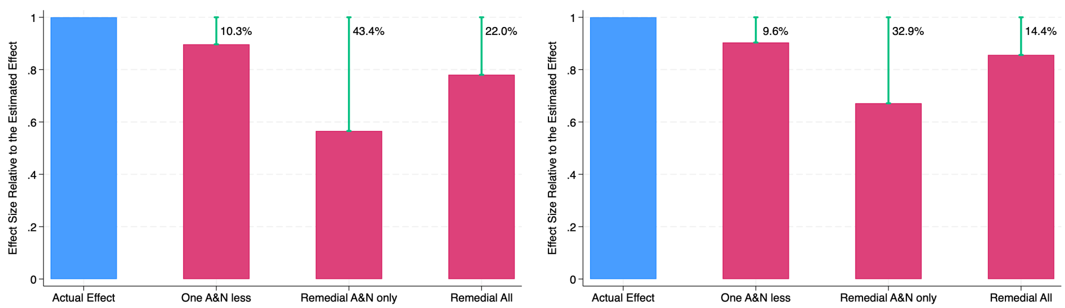


Figure 4: Total Score Losses due to A&N Students in Class

(a) Verbal (b) Numeric



Note: Data from Stockholm Birth Cohort Study. Figures present the score losses by the distance to an A&N peer according to the LIMM as estimated using 2SLS+CF estimator reported in Table 4 under different counterfactual scenarios. 1. *Actual Effect*, replicating our original results. 2. *One A&N less*, remove one A&N student from the classrooms. 3. *Remedial A&N only*, boosting scores of A&N students by 11 percent of a SD. 4. *Remedial All*, "subsidize" achievement of all 22 students so that each increases it by 0.5 percent of a SD. We simulated 1000 networks in which the dyad linking probabilities were such that the networks matched, on average, the number of edges and the clustering of the average classroom network we observe in the data. See mode details on the simulation in Online Appendix IV. Figure 3 presents the effects by distance to the closest A&N peer. Figure 4 presents the aggregate losses at the classroom level.

for verbal and numeric scores result from eliminating feedback loops where the two A&N peers were close to the same social groups. The largest relative effect reduction is in the counterfactual scenario where A&N students get remedial classes and boost their scores by 11 percent of a SD. It reduces the total effect by 43.4 percent in verbal and 32.9 percent in numeric ability. Figure 3 shows that most of the effect reductions achieved by the targeted policy come from the score losses avoided in the A&N students in particular. It also shows that the direct friends of the A&N student also benefit from the intervention, in particularly their verbal score. Finally, we show that subsidizing all students' scores by 0.5 percent of a SD reduces the total effect of having A&N peers by 22 percent in verbal and 14.4 percent in numeric scores. This policy shifts very little the effect for the affected peers, and it is, in fact, regressive, as those who face no negative impact from the A&N peers (i.e., those three friends away) see their scores improve.

6 Robustness

6.1 Missing links

The observed networks in our data are likely to suffer from mismeasurement. The number of links per student (i.e., edges per node) is top-coded in the sense that students can only nominate a maximum of three friends. We observe that roughly 60 percent of the students nominated three friends, suggesting that many would have nominated more had they had the chance to do so. Potential link censoring leads to the concern that distances to A&N peers might be shorter than observed and that some of the observed network components may be disjoint due to link censoring (Kossinets, 2006). Missing links due to top coding might lead us to incorrectly identify triads as intransitive when in reality, they are transitive. That may be problematic for identifying the structural parameters of the LIMM based on intransitive triads (Blume et al., 2015). Chandrasekhar and Lewis (2016) and Griffith (2022) show that censoring of networks either due to missing nodes or missing edges can lead to bias in the linear-in-means model and the reduced-form representation of it. While the two aforementioned studies offer constructive solutions to link censoring they are not particularly suitable for our application since they would require an uncensored subsample or at least certainty of uncensored nodes.[21]

In order to assess how sensitive our results are to network truncation we use the approach proposed in Patacchini et al. (2017) by artificially creating friendship links. First, we add one friend (i.e., a friendship nomination) to every student whose number of friends was top coded. We assign that link to the classmate with the closest **Z** vector who is currently not her friend. Second, we assign the link to a friend's friend who is currently not her friend. Third, we assign the link to a classmate who is currently not her friend but lives in the same neighborhood block. Finally, we add one friendship nomination to every student who nominated three friends *and* has an estimated degree heterogeneity above the 50th percentile. As with the first approach, we assign the link to the classmate with the closest **Z** vector who is currently not her friend.

Table V.2 in the Online Appendix tests how sensitive our distance results in Table 2 are to the network truncation issue by extending the observed networks following the four hypothetical scenarios listed above. It indicates that our results are robust to those changes in the networks. Similarly, in Table V.3 of Online Appendix V.2, we inquire how robust our structural estimates are to network truncation. The results remain remarkably robust.

In Online Appendix VII, we revisit the attrition of retained students discussed in Section 2. We probe its impact by running simulations in which we take the A&N children of the 1953 cohort that are not in the right grade for their age, randomly assign them to 1953-cohort classrooms, and synthetically link them to friends. We find that the negative relation between the size of the peer effect and distance to the A&N peer remains robust to recovering the attrited A&N retained students.

[21]Certainty of the absence of censoring cannot be established by dropping all top coded individuals since some 7 percent of the students in the classrooms did not participate in the sociometric survey. Nominations to non-participants are not observed.

6.2 Exclusion restriction of the link formation regression

As discussed in Section 3.2 and I, the identification hinges on the assumption that links are determined by a dyadic model (i.e., homophily variables are defined at the dyad-level for each i, j cell in the adjacency matrix) whereas the outcomes are determined at the individual level and hence the dyadic variables must be excluded from the outcome equation of individual i. Further, we construct our dyadic variables using arguably exogenous variables, namely prebirth and perinatal characteristics and family characteristics at birth. Nonetheless, how similar one is to one's classmates may both influence the likelihood of making friends and directly influence the outcome. We address this concern by running an informal test of the exclusion restriction that is based on the idea that the stability of the coefficients to observed characteristics is symptomatic of the stability to unobserved characteristics (Altonji et al., 2005; Hsieh and Kippersluis, 2018). We add to our main specification (8) $\frac{1}{n_r-1}\sum_{j\neq i\in r} z_{ij}$ for each $\mathbf{z}$ that yields binary homophily variables. That is the shares of classmates (or alternatively friends) who have the same value as i on the data that we use in the link formation model.[22] This way, we control for measures indicating how typical student i is relative to her classmates. The results reported in Table V.1 of Online Appendix V show remarkable stability relative to the benchmark estimates of our main specification in Table 2. These results indicate that (dis)similarities between classmates do not affect the peer effect.

6.3 Placebo treatment analyses

To test whether our estimates report spurious relationships, are driven by peculiarities in the data, or model mispecification, we implement placebo treatment analyses. We performed 1,000 iterations where we randomly assigned students the A&N status, recalculated our explanatory variables based on the new (fake) A&N peers, and estimated each of our models for the numeric and verbal scores. We then collect the

[22] Hsieh and Kippersluis (2018) control for aggregate (not average, as we do) differences between the individual and her friends in their local aggregate model (instead of a local average model).

estimated parameters and plot their distributions in Figures V.1 in online appendix V.4. They show that for both outcomes, the distribution of the estimated parameters are, as expected, tightly centered at zero. Furthermore, Figures V.1 show that our actual estimates are on the tails of these placebo-coefficient distributions. In particular, in the estimations regarding verbal ability, our estimated coefficients would be in the top-1 percent (Distance to A&N) or bottom 1 percent (A&N in class) of the placebo-coefficient distributions. Similarly, our estimated coefficients for numeric ability would be in the top-5 percent (Distance to A&N) or bottom 5 percent (A&N in class) of the placebo-coefficient distributions. These results indicate that the relationships we have identified between the location of the A&N peer and the size of the negative externality are genuine and not due to chance, modelling choices, or estimation flukes.

6.4 Falsification regressions

To further test whether it is our exogenous variation in the treatment (i.e., exposure to disruptive peers) that drives the results of our main specification reported in Table 2, we run a falsification test of our distance model (8) using the following predetermined placebo outcomes that could arguably not have been affected by one's A&N peers in sixth grade: the duration of mother's postpartum stay (days) after the cohort member's birth, height at birth, number of rooms in the apartment as of 1960, number of household members as of 1960, an indicator variable for own father having a white collar occupation. Although perhaps more tenuous than the aforementioned placebo outcomes, the cohort member's own height measured at enlistment (age 19) should arguably be resilient to abused and neglected peers. Table V.4 of Online Appendix V shows no evidence of distance to the A&N peers having an effect on the placebo outcomes. These null results lend further support to the treatment being the driver of our main estimates and not alternative processes related to selection or classroom manipulation by parents.

7 Conclusions

This study analyzes the diffusion of the negative externalities generated by parentally abused and neglected classroom peers. In the first part of the paper, we propose a parsimonious empirical model on the diffusion of classroom externalities, which captures the idea of a fade-out of the peer-effect as it progresses away from its source through the classroom's social network. Friends of the abused and neglected peer face substantially stronger adverse effects on cognitive outcomes than those who are more remotely connected to her. Our results show that being three friends (as measured by path length) away from the abused and neglected peer is equivalent to not having her in the classroom at all. These findings suggest that parentally abused and neglected peers do not uniformly impede everyone's learning and cognitive development. Rather, their experiences of abuse and neglect, externalizing behavior, and worse academic achievement imprint negative spillovers primarily on their closest friends. These spillovers dissipate quickly as the path length to the peer increases.

In the second part of the paper, we estimate the structural parameters of the LIMM using methods that help to overcome the reflection problem, thus providing a way to understand the nature of the peer effect. We find that, even though the students' losses in verbal and numeric abilities caused by abused and neglected peers have similar fade-out rates as path length increases, the nature of the effects through which those losses materialize differ. While abused and neglected students' disruptiveness affects peers' acquisition of verbal ability through the drop in the former's verbal performance, it depletes numeric skills directly.

Understanding the nature of the disruptive peer effects at work in the classroom has important implications for the design of policy responses. Our results suggest that while disruptiveness generally does not affect the learning process of every classmate (e.g., stopping class, diverting resources to the disruptive peer), it does significantly harm the cognitive acquisition of the disruptive peer as well as her social circle. Our counterfactual policy experiments show that interventions aimed at reducing the negative consequences of parentally abused and neglected peers should therefore specifically target the student in question and her immediate friends. Class-wide interventions are accordingly less efficient, as some resources would be devoted to students who are not affected by the disruptiveness.

Appendix

Table A.1: The Association between Abuse and Neglect and Own Achievement and Behavioral Outcomes during Adolescence

Cognitive measures & educational attainment						
	IQ Components at 13			Years of	Drop out of	Any post
	Verbal	Numeric	Spatial	education	high school	secondary
Abused &	-2.361	-2.498	-1.468	-0.667	0.147	-0.086
Neglected	(0.243)	(0.301)	(0.267)	(0.081)	(0.019)	(0.013)
Observations	11,434	11,444	11,431	11,476	11,476	11,476

Acting out during adolescence (Child Welfare Committee reports)						
	Adjustment problems†	Drugs	Violence	Vandalism	Drinking	Delinquency
Abused &	0.084	0.063	0.035	0.011	0.049	0.108
Neglected	(0.012)	(0.010)	(0.009)	(0.006)	(0.010)	(0.014)
Observations	12,047	12,047	12,047	12,047	12,047	12,047

Note: Data from Stockholm Birth Cohort Study. † Adjustment problems refers to adjustment problems at home (e.g., running away), at school (truancy), or attempted suicide. The regression includes school fixed effects and gender controls. Robust standard errors in parentheses.

Table A.2: Exposure to Abused and Neglected Peers and Own Parental Abuse and Neglect

	Dependent variable: Own parental abuse & neglect	
	Coefficient	*Standard error*
Percent abused & neglected male peers	-0.019	(0.116)
Percent abused & neglected female peers	0.067	(0.104)
Observations	7,995	

Note: Data from Stockholm Birth Cohort Study. Sample excludes classes with less than seven students, schools with only one class in grade 6, schools with special education classrooms and classrooms in which no one participated in the sociometric survey. Robust standard errors in parentheses clustered at the school level. Following Guryan et al. (2009).

Table A.3: Balancing Test with Continuous Exposure: Fraction of Disruptive Peers and Own Characteristics

	Gender (1)	Gender (2)	Social aid receipt (1)	Social aid receipt (2)	Birth weight (1)	Birth weight (2)	Mother's age at birth (1)	Mother's age at birth (2)
Share of A&N peers in class								
All	-0.001	-0.001	0.029	0.002	-0.036	-0.009	0.005	0.001
	(0.002)	(0.001)	(0.007)	(0.003)	(0.018)	(0.014)	(0.003)	(0.001)
Male	0.001	0.000	0.014	0.001	-0.008	0.000	0.002	0.000
	(0.001)	(0.001)	(0.003)	(0.002)	(0.012)	(0.010)	(0.001)	(0.001)
Female	-0.002	-0.001	0.016	0.001	-0.028	-0.009	0.003	0.000
	(0.001)	(0.001)	(0.004)	(0.002)	(0.011)	(0.009)	(0.002)	(0.001)
Shoool FE	N	Y	N	Y	N	Y	N	Y
Observations	7,549	7,549	7,549	7,549	6,091	6,091	6,255	6,255

	Owner of dwelling (1)	Owner of dwelling (2)	Dwelling size (1)	Dwelling size (2)	Older siblings (1)	Older siblings (2)
Share of A&N peers in class						
All	-0.020	0.004	0.034	0.029	0.000	-0.001
	(0.004)	(0.002)	(0.003)	(0.003)	(0.001)	(0.001)
Male	-0.008	0.002	0.018	0.015	-0.001	-0.001
	(0.003)	(0.001)	(0.002)	(0.002)	(0.001)	(0.000)
Female	-0.012	0.002	0.017	0.015	0.000	0.000
	(0.003)	(0.001)	(0.002)	(0.002)	(0.001)	(0.000)
Shoool FE	N	Y	N	Y	N	Y
Observations	7,549	7,549	7,549	7,549	7,549	7,549

Note: Data from Stockholm Birth Cohort Study. Each entry reports the slope coefficient from a separate regression. Sample excludes classes with less than seven students, schools with only one class in grade 6, schools with special education classrooms and classrooms in which no one participated in the sociometric survey. Sample restricted to individuals who were not themselves abused and neglected (Angrist, 2014). Robust standard errors in parentheses clustered at the school level.

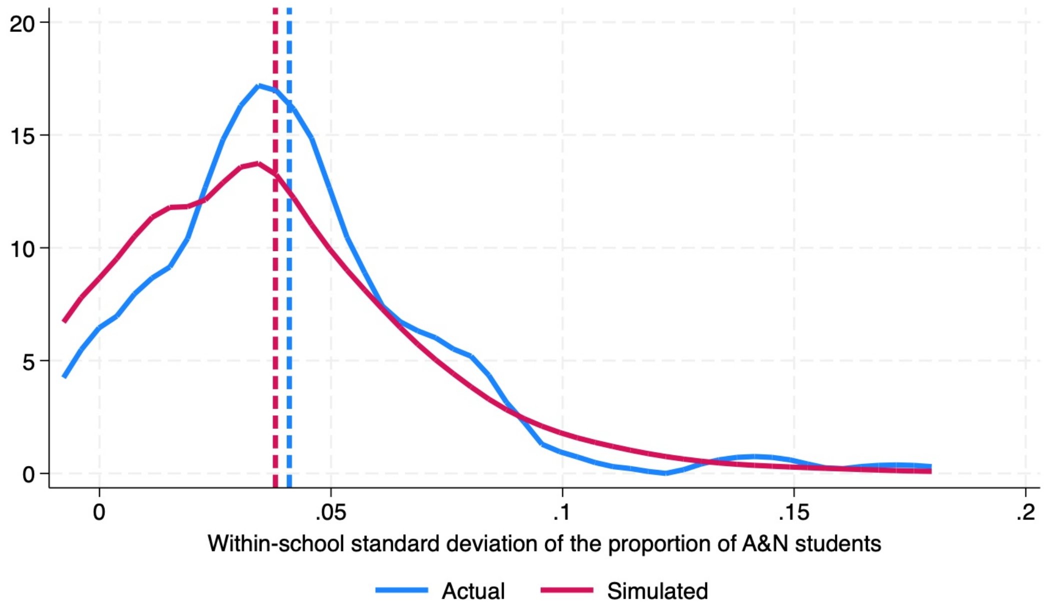
Table A.4: Balancing Test with Binary Exposure: Exposure to at Least One Disruptive Peers and Own Characteristics

	Gender		Social aid receipt		Birth weight		Mother's age at birth	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
At least one A&N peer in class								
All	-0.018	-0.009	0.158	-0.001	-0.224	-0.083	0.006	-0.009
	(0.014)	(0.009)	(0.027)	(0.015)	(0.125)	(0.109)	(0.016)	(0.010)
Male	0.002	0.000	0.135	0.004	-0.056	-0.020	0.009	-0.001
	(0.014)	(0.011)	(0.035)	(0.016)	(0.135)	(0.120)	(0.019)	(0.011)
Female	-0.022	-0.016	0.143	-0.005	-0.348	-0.129	0.018	0.001
	(0.013)	(0.008)	(0.028)	(0.018)	(0.122)	(0.097)	(0.016)	(0.009)
Shoool FE	N	Y	N	Y	N	Y	N	Y
Observations	7,549	7,549	7,549	7,549	6,091	6,091	6,255	6,255

	Owner of dwelling		Dwelling size		Older siblings	
	(1)	(2)	(1)	(2)	(1)	(2)
At least one A&N peer in class						
All	-0.113	0.056	0.384	0.319	0.000	0.003
	(0.041)	(0.024)	(0.032)	(0.031)	(0.007)	(0.005)
Male	-0.081	0.025	0.279	0.224	-0.006	-0.007
	(0.039)	(0.022)	(0.028)	(0.024)	(0.007)	(0.005)
Female	-0.100	0.049	0.245	0.199	0.002	0.007
	(0.037)	(0.021)	(0.028)	(0.029)	(0.007)	(0.005)
Shoool FE	N	Y	N	Y	N	Y
Observations	7,549	7,549	7,549	7,549	7,549	7,549

Note: Data from Stockholm Birth Cohort Study. Each entry reports the slope coefficient from a separate regression. Sample excludes classes with less than seven students, schools with only one class in grade 6, schools with special education classrooms and classrooms in which no one participated in the sociometric survey. Sample restricted to individuals who were not themselves abused and neglected (Angrist, 2014). Robust standard errors in parentheses clustered at the school level.

Figure A.1: Actual and Simulated Within-school Standard Deviation in the Proportion A&N Students



Note: The figure shows the observed and simulated distributions of the within-school standard deviations of the classroom proportion of A&N students as in Lavy et al. (2011). We obtain the simulated distribution by randomly generating the A&N status of the students in each school using a binomial distribution with a p equal to the proportion of A&N students in the school. Using the simulated data, we calculate the within-school standard deviation of the classroom proportion of A&N students. We repeat this process 1,000 times. Vertical lines indicate the mean of each distribution.

References

Aizer, Anna (2008), "Peer Effects and Human Capital Accumulation: the Externalities of ADD." NBER Working Paper. 1, 3

Almlund, Mathilde, Angela Lee Duckworth, James Heckman, and Tim Kautz (2011), "Personality Psychology and Economics." *Handbook of the Economics of Education*, 4, 1–181, URL http://dx.doi.org/10.1016/B978-0-444-53444-6.00001-8. 2

Almquist, Y. B. (2014). (2014), "Stockholm Birth Cohort. Version 1." https://researchdata.se/en/catalogue/dataset/ext0106-1. 2

Altonji, Joseph G, Todd E Elder, and Christopher R Taber (2005), "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy*, 113 (1), 151–184, URL https://www.journals.uchicago.edu/doi/10.1086/426036. 6.2

Angrist, Joshua (2014), "The Perils of Peer Effects." *Labour Economics*, 30, 98–108, URL http://linkinghub.elsevier.com/retrieve/pii/B9780444534446000018. 3.2.4, 13, 2, A.3, A.4, II.1

Arduini, Tiziano, Eleonora Patacchini, and Eduardo Rainone (2015), "Parametric and Semiparametric IV Estimation of Network Models with Selectivity." Technical report, Einaudi Institute for Economics and Finance (EIEF). 1

Attanasio, Orazio, Abigail Barr, Juan Camilo Cardenas, Garance Genicot, and Costas Meghir (2012), "Risk Pooling, Risk Preferences, and Social Networks." *American Economic Journal: Applied Economics*, 4 (2), 134–167, URL http://pubs.aeaweb.org/doi/10.1257/app.4.2.134. 3.2.3

Auerbach, Eric (2022a), "Identification and Estimation of a Partially Linear Regression Model Using Network Data." *Econometrica*, 90 (1), 347–365, URL https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA19794. 1

Auerbach, Eric (2022b), "Identification and Estimation of a Partially Linear Regression Model Using Network Data." *Econometrica*, 90 (1), 347–365, URL https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA19794. 1, I

Badev, Anton (2021), "Nash Equilibria on (Un)Stable Networks." *Econometrica*, 89 (3), 1179–1206, URL https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA12576. 12

Banerjee, A, A G Chandrasekhar, E Duflo, and M O Jackson (2013), "The Diffusion of Microfinance." *Science*, 341 (6144), 1236498–1236498, URL http://www.sciencemag.org/cgi/doi/10.1126/science.1236498. 1

Blume, Lawrence, William Brock, Steven Durlauf, and Yannis Ioannides (2011), "Identification of Social Interactions." *Handbook of Social Economics*, 854–965. 3.1.1

Blume, Lawrence E, William A Brock, Steven N Durlauf, and Rajshri Jayaraman (2015), "Linear Social Interactions Models." *Journal of Political Economy*, 123 (2), 444–496, URL http://www.journals.uchicago.edu/doi/10.1086/679496. 6.1

Borghans, Lex, Angela Lee Duckworth, James J Heckman, and Bas Ter Weel (2008), "The Economics and Psychology of Personality Traits." *Journal of Human Resources*, 43 (4), 972–1059, URL http://muse.jhu.edu/login?auth=0&type=summary&url=/journals/journal_of_human_resources/v043/43.4.borghans.html. 2

Boucher, V. and E. A. Houndetoungan (2023), "Estimating Peer Effects using Partial Network Data." Unpublished. 2, VII

Boucher, Vincent, Carlo L. Del Bello, Fabrizio Panebianco, Thierry Verdier, and Yves Zenou (2022), "Education Transmission and Network Formation." *Journal of Labor Economics*, 0 (0), –, URL https://doi.org/10.1086/718981. 3.2.3

Boucher, Vincent and Bernard Fortin (2016), "Some Challenges in the Empirics of the Effects of Networks." In *The Oxford Handbook of Economics of Networks* (Yann Bramoulle, Andrea Galeotti, and Brian W. Rogers, eds.), chapter 10, 277–302, Oxford University Press, Oxford. 19

Boucher, Vincent, Michelle Rendall, Philip Ushchev, and Yves Zenou (2024), "Toward a general theory of peer effects." *Econometrica*, 92 (2), 543–565, URL https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA21048. 3.1.1

Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin (2009), "Identification of Peer Effects Through Social Networks." *Journal of Econometrics*, 150 (1), 41–55, URL http://linkinghub.elsevier.com/retrieve/pii/S0304407609000335. 1, 3.1, 3.1.1, 5, 5.1.1, 5.1.1

Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin (2020), "Peer Effects in Networks: A Survey." *Annual Review of Economics*, 12 (1), 603–629, URL https://doi.org/10.1146/annurev-economics-020320-033926. 19

Brock, William and Steven Durlauf (2001), "Discrete Choice with Social Interactions." *The Review of Economic Studies*, 68 (2), 235–260. 3.1.1

Calvó-Armengol, Antoni, Eleonora Patacchini, and Yves Zenou (2009), "Peer Effects and Social Networks in Education." *The Review of Economic Studies*, 76 (4), 1239–1267, URL http://restud.oxfordjournals.org/content/76/4/1239.full. 1, 3.1.2

Carlson, Bonnie E (2000), "Children Exposed to Intimate Partner Violence Research Findings and Implications for Intervention." *Trauma, Violence, & Abuse*, 1 (4), 321–342, URL http://tva.sagepub.com/content/1/4/321.abstract. 2

Carrell, S E, B I Sacerdote, and J E West (2013), "From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation." *Econometrica*, 81 (3), 855–882, URL http://onlinelibrary.wiley.com/doi/10.3982/ECTA10168/full. 1, 3.2.3

Carrell, Scott E and Mark L Hoekstra (2010), "Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone's Kids." *American Economic Journal: Applied Economics*, 2 (1), 211–228, URL http://pubs.aeaweb.org/doi/abs/10.1257/app.2.1.211. 1, 2, 3, 3.2.1, 15, II

Carrell, Scott E, Mark L Hoekstra, and Elira Kuka (2018), "The Long-Run Effects of Disruptive Peers." *American Economic Review*, 108, 3377–3415. 1, 3, 15

Chandrasekhar, Arun and Randall Lewis (2016), "Econometrics of Sampled Networks." https://stanford.edu/~arungc/CL.pdf. 6.1

Coleman, James Samuel (1964), *Introduction to Mathematical Sociology*. Free Press, URL http://books.google.com/books?id=T7NEAAAAIAAJ&q=Introduction+to+mathematical+sociology&dq=Introduction+to+mathematical+sociology&hl=&cd=2&source=gbs_api. 3.1.1

Dahl, Gordon B, Katrine V Løken, and Magne Mogstad (2014), "Peer Effects in Program Participation." *American Economic Review*, 104 (7), 2049–2074, URL http://pubs.aeaweb.org/doi/10.1257/aer.104.7.2049. 1

de Chaisemartin, Clément and Nicolás Navarrete H. (2023), "The Direct and Spillover Effects of a Nationwide Socioemotional Learning Program for Disruptive Students." *Journal of Labor Economics*, 41 (3), 729–769, URL https://doi.org/10.1086/720455. 3

De Giorgi, Giacomo, Michele Pellizzari, and Silvia Redaelli (2010), "Identification of Social Interactions through Partially Overlapping Peer Groups." *American Economic Journal: Applied Economics*, 2 (2), 241–275, URL http://pubs.aeaweb.org/doi/10.1257/app.2.2.241. 5, 5.1.1

de Paula, Áureo (2017), "Econometrics of Network Models." In *Advances in Economics and Econometrics* (Bo Honore, Ariel Pakes, Monika Piazzesi, and Larry Samuelson, eds.), 268–323, Cambridge University Press, Cambridge, URL https://www.cambridge.org/core/product/identifier/CBO9781108227162A017/type/book_part. 1, 3.1.1

de Paula, Aureo (2020), "Chapter 3 - Strategic Network Formation." In *The Econometric Analysis of Network Data* (Bryan Graham and Aureo de Paula, eds.), 41–61, Academic Press, URL https://www.sciencedirect.com/science/article/pii/B9780128117712000092. 3.2.3

Díaz, Carlos, Eleonora Patacchini, Thierry Verdier, and Yves Zenou (2021), "Leaders in Juvenile Crime." *Journal of Economic Behavior & Organization*, 192, 638–667, URL https://linkinghub.elsevier.com/retrieve/pii/S0167268121004649. 1, 5.3

Eriksen, Tine Louise Mundbjerg, Helena Skyt Nielsen, and Marianne Simonsen (2014), "Bullying in Elementary School." *Journal of Human Resources*, 49 (4), 839–871, URL http://jhr.uwpress.org/lookup/doi/10.3368/jhr.49.4.839. 2

Figlio, David (2005), "Boys Named Sue: Disruptive Children and Their Peers." NBER Working Paper. 3

Fredriksson, Peter, Bjorn Öckert, and Hessel Oosterbeek (2012), "Long-Term Effects of Class Size." *The Quarterly Journal of Economics*, 128 (1), 249–285, URL https://academic.oup.com/qje/article-lookup/doi/10.1093/qje/qjs048. 11

Gazze, Ludovica, Claudia L. Persico, and Sandra Spirovska (2022), "The Long-Run Spillover Effects of Pollution: How Exposure to Lead Affects Everyone in the Classroom." *Journal of Labor Economics*, 723175, URL https://www.journals.uchicago.edu/doi/10.1086/723175. 3

Goldsmith-Pinkham, Paul and Guido W Imbens (2013), "Social Networks and the Identification of Peer Effects." *Journal of Business & Economic Statistics*, 31 (3), 253–264, URL http://www.tandfonline.com/doi/abs/10.1080/07350015.2013.801251. 1, 3.1.1, I

Gould, Eric D, Victor Lavy, and M Daniele Paserman (2009), "Does Immigration Affect the Long-Term Educational Outcomes of Natives? Quasi-Experimental Evidence." *The Economic Journal*, 119 (540), 1243–1269, URL http://doi.wiley.com/10.1111/j.1468-0297.2009.02271.x. 3

Graham, Bryan S (2015), "Methods of Identification in Social Networks." *Annual Review of Economics*, 7 (1), 465–485, URL http://www.annualreviews.org/doi/10.1146/annurev-economics-080614-115611. 12

Graham, Bryan S (2017), "An Econometric Model of Network Formation With Degree Heterogeneity." *Econometrica*, 85 (4), 1033–1063, URL https://www.econometricsociety.org/doi/10.3982/ECTA12679. 1, 3.2.3, 12, I

Graham, Bryan S (2018), "Identifying and Estimating Neighborhood Effects." *Journal of Economic Literature*, 56 (2), 450–500. 1

Graham, Bryan S. (2020), "Chapter 2 - Dyadic Regression." In *The Econometric Analysis of Network Data* (Bryan Graham and Aureo de Paula, eds.), 23–40, Academic Press, URL https://www.sciencedirect.com/science/article/pii/B9780128117712000080. 3.2.3

Griffith, Alan (2022), "Name Your Friends, but Only Five? The Importance of Censoring in Peer Effects Estimates Using Social Network Data." *Journal of Labor Economics*, 40 (4), 779–805, URL https://doi.org/10.1086/717935. 6.1

Guryan, Jonathan, Kory Kroft, and Matthew J Notowidigdo (2009), "Peer Effects in the Workplace: Evidence from Random Groupings in Professional Golf Tournaments." *American Economic Journal: Applied Economics*, 1 (4), 34–68, URL http://pubs.aeaweb.org/doi/10.1257/app.1.4.34. 3.2.1, A.2

Hart, Betty and Todd Risley (1995), *Meaningful Differences in the Everyday Experience of Young American Children*. Paul H. Brookes Publishing Co., Inc, Baltimore, MD. 5.2

Holt, Stephanie, Helen Buckley, and Sadhbh Whelan (2008), "The Impact of Exposure to Domestic Violence on Children and Young People: A Review of the Literature." *Child Abuse & Neglect*, 32 (8), 797–810, URL http://linkinghub.elsevier.com/retrieve/pii/S0145213408001348. 2

Hoxby, C (2000), "Peer Effects in the Classroom: Learning from Race and Gender." NBER Working paper. 1, 3.2.2

Hsieh, Chih-Sheng and Hans Kippersluis (2018), "Smoking Initiation: Peers and Personality." *Quantitative Economics*, 9 (2), 825–863–863, URL http://qeconomics.org/ojs/index.php/qe/article/view/609. 6.2, 22

Hsieh, Chih-Sheng, Michael D. Konig, and Xiaodong Liu (2022), "A Structural Model for the Coevolution of Networks and Behavior." *The Review of Economics and Statistics*, 104 (2), 355–367. 1

Hsieh, Chih-Sheng and Lung-Fei Lee (2015), "A Social Interactions Model with Endogenous Friendship Formation and Selectivity." *Journal of Applied Econometrics*, 31 (2), 301–319, URL http://doi.wiley.com/10.1002/jae.2426. 1, 3.2.3, I

Hsieh, Chih-Sheng, Lung-Fei Lee, and Vincent Boucher (2020), "Specification and Estimation of Network Formation and Network Interaction Models with the Exponential Probability Distribution." *Quantitative Economics*, 11 (4), 1349–1390, URL https://onlinelibrary.wiley.com/doi/abs/10.3982/QE944. 1

Husen, Torsten (1961), "School Reform in Sweden: A Liberal Democracy Adopts the Comprehensive School System." *The Phi Delta Kappan*, 43 (2), 86–91, URL http://www.jstor.org/stable/20342681. 3.2.2

Jackson, Matthew O (2010), *Social and Economic Networks*. Princeton University Press, URL http://books.google.com/books?id=rFzHinVAq7gC&printsec=frontcover&dq=social+and+economic+networks&hl=&cd=1&source=gbs_api. 3.1.1, 3.2.3

Jackson, Matthew O and Brian W Rogers (2007), "Meeting Strangers and Friends of Friends: How Random Are Social Networks?" *American Economic Review*, 97 (3), 890–915, URL http://pubs.aeaweb.org/doi/10.1257/aer.97.3.890. 3.1.2, 3.1.2, 3.2.3

Jackson, Matthew O and Asher Wolinsky (1996), "A Strategic Model of Social and Economic Networks." *Journal of Economic Theory*, 71 (1), 44–74, URL http://linkinghub.elsevier.com/retrieve/pii/S0022053196901088. 1, 12

Jackson, Matthew O. and Leeat Yariv (2011), "Chapter 14 - Diffusion, Strategic Interaction, and Social Structure." volume 1 of *Handbook of Social Economics*, 645–678, North-Holland, URL https://www.sciencedirect.com/science/article/pii/B9780444531872000140. 1

Johnsson, Ida and Hyungsik Roger Moon (2021), "Estimation of Peer Effects in Endogenous Social Networks: Control Function Approach." *The Review of Economics and Statistics*, 103 (2), 328–345, URL https://doi.org/10.1162/rest_a_00870. 1, 3.2.3, 5, 5.1.2, 17, I, V.3

Kelejian, Harry H and Gianfranco Piras (2014), "Estimation of Spatial Models with Endogenous Weighting Matrices, and an Application to a Demand Model for Cigarettes." *Regional Science and Urban Economics*, 46 (C), 140–149, URL http://dx.doi.org/10.1016/j.regsciurbeco.2014.03.001. 14

König, Michael D., Xiaodong Liu, and Yves Zenou (2019), "R&D Networks: Theory, Empirics, and Policy Implications." *The Review of Economics and Statistics*, 101 (3), 476–491, URL https://doi.org/10.1162/rest_a_00762. 14

Kossinets, Gueorgi (2006), "Effects of Missing Data in Social Networks." *Social Networks*, 28 (3), 247–268, URL https://linkinghub.elsevier.com/retrieve/pii/S0378873305000511. 6.1

Lavy, Victor, M Daniele Paserman, and Analia Schlosser (2011), "Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers in the Classroom." *The Economic Journal*, 122 (559), 208–237, URL http://doi.wiley.com/10.1111/j.1468-0297.2011.02463.x. 3, 3.2.2, A.1

Lin, Xu (2010), "Identifying Peer Effects in Student Academic Achievement by Spatial Autoregressive Models with Group Unobservables." *Journal of Labor Economics*, 28 (4), 825–860, URL http://www.journals.uchicago.edu/doi/10.1086/653506. 3.2.3

List, John A, Fatemeh Momeni, and Yves Zenou (2020), "The Social Side of Early Human Capital Formation: Using a Field Experiment to Estimate the Causal Impact of Neighborhoods." Working Paper 28283, National Bureau of Economic Research, URL http://www.nber.org/papers/w28283. 1

Liu, Xiaodong, Eleonora Patacchini, and Yves Zenou (2014), "Endogenous Peer Effects: Local Aggregate or Local Average?" *Journal of Economic Behavior & Organization*, 103, 39–59, URL https://linkinghub.elsevier.com/retrieve/pii/S0167268114001036. 3.1.1

Manski, Charles (1993), "Identification of Endogenous Social Effects: The Reflection Problem." *The Review of Economic Studies*, 60 (3), 531–542. 1, 3.1, 3.1.1, 10, 5

Manski, Charles F (2013a), "Comment." *Journal of Business & Economic Statistics*, 31 (3), 273–275, URL http://www.tandfonline.com/doi/abs/10.1080/07350015.2013.792262. 1

Manski, Charles F. (2013b), "Identification of Treatment Response with Social Interactions." *The Econometrics Journal*, 16 (1), S1–S23, URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1368-423X.2012.00368.x. 1

McPherson, Miller, Lynn Smith-Lovin, and James M Cook (2001), "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology*, 27, 415–444. 3.2.3

Meghir, Costas, Mårten Palme, and Emilia Simeonova (2013), "Education, Cognition and Health: Evidence from a Social Experiment." Working Paper 19002, National Bureau of Economic Research, URL http://www.nber.org/papers/w19002. 20

Moylan, Carrie A, Todd I Herrenkohl, Cindy Sousa, Emiko A Tajima, Roy C Herrenkohl, and M Jean Russo (2009), "The Effects of Child Abuse and Exposure to Domestic Violence on Adolescent Internalizing and Externalizing Behavior Problems." *Journal of Family Violence*, 25 (1), 53–63, URL http://link.springer.com/10.1007/s10896-009-9269-9. 2

Opper, Isaac M (2019), "Does Helping John Help Sue? Evidence of Spillovers in Education." *American Economic Review*, 109 (3), 1080–1115, URL https://pubs.aeaweb.org/doi/10.1257/aer.20161226. 1

Özek, Umut (2021), "The Effects of Middle School Remediation on Postsecondary Success: Regression Discontinuity Evidence from Florida." *Journal of Public Economics*, 203, 104518, URL https://linkinghub.elsevier.com/retrieve/pii/S0047272721001547. 5.3

Patacchini, Eleonora, Edoardo Rainone, and Yves Zenou (2017), "Heterogeneous Peer Effects in Education." *Journal of Economic Behavior & Organization*, 134, 190–227, URL http://dx.doi.org/10.1016/j.jebo.2016.10.020. 4, 3.2.3, 6.1, I

Patacchini, Elonora and Yves Zenou (2012), "Juvenile Delinquency and Conformism." *Journal of Law, Economics, and Organization*, 28 (1), 1–31, URL https://academic.oup.com/jleo/article-lookup/doi/10.1093/jleo/ewp038. 3.1.1

Paulston, Rolland G. (1966), "The Swedish Comprehensive School Reform: A Selected Annotated Bibliography." *Comparative Education Review*, 10 (1), 87–94, URL http://www.jstor.org/stable/1186319. 3.2.2

Purpura, David J., Sara A. Schmitt, and Colleen M. Ganley (2017), "Foundations of Mathematics and Literacy: The Role of Executive Functioning Components." *Journal of Experimental Child Psychology*, 153, 15–34, URL https://www.sciencedirect.com/science/article/pii/S0022096516301357. 5.2

Sacerdote, Bruce (2011), "Chapter 4 - Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?" volume 3 of *Handbook of the Economics of Education*, 249–277, Elsevier, URL https://www.sciencedirect.com/science/article/pii/B9780444534293000041. 1, 1

Santavirta, Torsten and Miguel Sarzosa (2025), "Replication Package for: "Disruptive Peers in Endogenous Social Networks"." Data and code on Zenodo. 5, 2

Sarzosa, Miguel (2021), "Victimization and Skill Accumulation: The Case of School Bullying." *Journal of Human Resources*, URL https://jhr.uwpress.org/content/early/2021/11/03/jhr.0819-10371R2. 3, 3.2.1

Sarzosa, Miguel and Sergio Urzua (2021), "Bullying Among Adolescents: The Role of Skills." *Quantitative Economics*, 12 (3). 2, 3

Schennach, Susanne M. (2018), "Long memory via networking." *Econometrica*, 86 (6), 2221–2248, URL https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA11930. 1

Sheng, Shuyang (2020), "A Structural Econometric Analysis of Network Formation Games Through Subnetworks,." *Econometrica*, 88 (5), 1829–1858, URL https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA12558. 12

SOU1961:30 (1961), *Grundskolan. Betänkande avgivet av 1957 års skolberedning*, volume (The Comprehensive School. Report by the 1957 School Committee). Ivar Häggströms Boktryckeri AB, Stockholm, URL https://lagen.nu/sou/1961:30. 3.2.2

SOU1974:7 (1974), "Barn och ungdomsvård: Ett betänkande av fosterbarnsutredningen [Engl. Child Protection: A Report of the Foster Care Inquiry]." 2

Statistics Sweden (2023), "Statistik om Socialtjänstinsatser till Barn och Unga 2022 [Engl. Child Welfare Statistics 2022]." 2

Stehlé, Juliette, François Charbonnier, Tristan Picard, Ciro Cattuto, and Alain Barrat (2013), "Gender Homophily from Spatial Behavior in a Primary School: A Sociometric Study." *Social Networks*, 35, 604–613. 5.3

Stenberg, Sten-Åke and Denny Vågerö (2006), "Cohort Profile: The Stockholm Birth Cohort of 1953." *International Journal of Epidemiology*, 35 (3), 546–548, URL https://academic.oup.com/ije/article-lookup/doi/10.1093/ije/dyi310. 5

Svensson, Allan (1964), *Sociala och Regionala Faktorers Samband med Över- och Underrepresentation i Skolarbetet [Engl. Social and Regional Factors Explaining Under and Overrepresentation at School]*. (Mimeographed). 2

Svensson, Allan (1971), *Relative Achievement. School Performance in Relation to Intelligence, Sex and Home Environment*. Göteborg Studies in Educational Sciences, Almqvist & Wiksell, Stockholm. 2

Topa, Giorgio and Yves Zenou (2014), "Neighborhood and Network Effects." In *Handbook of Regional and Urban Economics* (Gilles Duranton, Vernon Henderson, and William Strange, eds.), volume 5, chapter 9, North-Holland, Amsterdam, The Netherlands. 3.1.1

Ushchev, Philip and Yves Zenou (2020), "Social Norms in Networks." *Journal of Economic Theory*, 185, 104969, URL https://www.sciencedirect.com/science/article/pii/S0022053119301206. 3.1.1

Wolfe, David A, Claire V Crooks, Vivien Lee, Alexandra McIntyre-Smith, and Peter G Jaffe (2003), "The Effects of Children's Exposure to Domestic Violence: A Meta-Analysis and Critique." *Clinical Child and Family Psychology Review*, 6 (3), 171–187. 2

Online Appendix for Effects of Disruptive Peers in Endogenous Social Networks

Torsten Santavirta Miguel Sarzosa

For online publication: Online appendix

I Link formation regression results

Table I.1: Summary Statistics of the Homophily Variables

Variable	Mean	Std. Dev.	Observations
Female	0.512	0.500	7,995
Parents social aid recipients	0.118	0.322	7,995
Owner of dwelling	0.181	0.385	7,995
Death of a sibling (pregnancy or postpartum stay)	0.160	0.367	6,480
First pregnancy	0.387	0.487	6,480
Prenatal care by doctor	0.425	0.494	6,470
Natal health 1	-0.018	1.399	6,275
Natal health 2	-0.006	1.297	6,275
Father's Occupational Status in 1953			
Upper & upper middle class	0.146	0.354	7,777
Middle class	0.336	0.472	7,777
Lower middle class	0.064	0.245	7,777
Skilled workers	0.280	0.449	7,777
Unskilled workers	0.173	0.378	7,777

Note: Data from Stockholm Birth Cohort Study. Sample excludes classes with less than seven students, schools with only one class in grade 6, schools with special education classrooms and classrooms in which no one participated in the sociometric survey. *Natal Health 1* and *Natal Health 2* are health endowment indices that consolidate information on prenatal and perinatal health of the child that comes from birth records in 1953. They are the first two principal components of a system of measures that include: mother's age at birth, hospital delivery, length of postpartum hospital stay, pre-eclampsia during pregnancy, other health conditions during pregnancy, anemia during pregnancy, fever during delivery, C-section, whether the delivery was facilitated by forceps, birthweight, and length at birth. *Natal Health 1* and *Natal Health 2* collect 14 percent and 12.3 percent of the measurement system variation respectively.

We estimate the link formation equation (9) of the main text with a joint maximum likelihood estimator proposed by Graham (2017) that identifies the unobserved heterogeneity of the sender i and receiver j, jointly. The observed dyadic covariate space Z_{ij} contains variables indicating whether i and j have the same gender, have a mother who received prenatal care by a physician during pregnancy, their fathers had the same SES in 1953, they lived in the same neighborhood block in 1953[23], both families owned their dwelling, and both households received social aid between 1953-1959. Z_{ij} also contains the dyadic distances between i and j in two health endowment indices that summarize their prenatal and perinatal health based on information from the birth records (i.e., Neonatal health 1 and Neonatal health 2).[24] Modeling link formation between individuals i and j based on the utility derived from similarity in observed and unobserved characteristics is common identification strategy to deal with endogeneity in peer effect studies (Goldsmith-Pinkham and Imbens, 2013; Hsieh and Lee, 2015; Graham, 2017; Patacchini et al., 2017; Johnsson and Moon, 2021; Auerbach, 2022b). The predetermined dyadic-level regressors are arguably excluded at the dyad-level from the outcome equation. For example, the distance in health endowments between student i and j does affect the probability of forming a link ($d_{ij} = 1$) (as seen in Table I.2) but at the dyad-level, it does not affect i's of j's outcomes. Hence, the identification comes from exclusion restrictions through functional form (Hsieh and Lee, 2015).

[23]We use the finest level of neighborhood available in SBC for year 1953 (cf. Codebook III). For the individuals who were born outside of Stockholm county but still belonged to the sampling frame (i.e., in-movers who lived in Stockholm as of 1963), we replaced neighborhood in 1953 with neighborhood in 1963. In total, the students in our analytic sample lived in 375 neighborhoods.

[24]See the descriptive statistics of the characteristics included in Z_{ij} as dyad-specific variables in Table I.1 in the Appendix.

Table I.2: Probability of Link Formation Between Two Classmates

Dependent variable:	Link btw i and j: d_{ij}
Homophily in terms of similarity: $\mathbf{1}[Z_i = Z_j]$	
Death of a sibling (pregnancy or postpartum stay)	0.001
	(0.02)
First born	0.096
	(3.06)
Prenatal care by doctor	0.043
	(1.36)
Same gender	4.097
	(85.56)
Social aid receipients	0.317
	(6.11)
Owner of dwelling	0.196
	(4.49)
Father's occ status, 1953	0.044
	(1.52)
Neighborhood	0.278
	(4.86)
Homophily in terms of distance: $\lvert Z_i - Z_j \rvert$	
Natal health index 1	0.003
	(0.90)
Natal health index 2	-0.008
	(-2.15)
Observations	92,287

Note: Data from Stockholm Birth Cohort Study. Sample excludes classes with less than seven students, schools with only one class in grade 6, schools with special education classrooms and classrooms in which no one participated in the sociometric survey. We further exclude classrooms in which no one participated in the sociometric survey. We model undirected links and estimate a fixed-effect logit regression where the dependent variable d_{ij} takes on value one if i nominated j or j nominated i, i.e., $d_{ji} = 1$ (and zero otherwise). All variables in the top panel are indicator variables taking on value 1 if i and j take on the same value on the underlying binary dummy Z. Homophily would imply positive signs of the coefficients of these sameness variables. The two variables in the bottom panel measure the difference in the continuous values of the underlying covariates Z. Homophily would imply negative signs of the coefficients of these distance variables. The number of observations (i.e., potential undirected links) in the dyadic-level data used in the regression contains 92,287 observations that stem from a sample of 7,995 unique students. The model includes joint fixed effects (sender i and receiver j). *Address53* is the individual's neighborhood of residence at birth at the precision of block in many cases. The students in the analytic sample lived in 375 different neighborhoods in 1953. Dummies to control for missing values at the students level are included. Their coefficients are provided upon request. t-statistics in parentheses.

II The presence of a social effect of disruptive peers

In this appendix, we empirically document the presence of a social effect in our context of spillovers of disruptive peers. Here, we estimate the average total effect using the reduced form of the LIMM. The scholastic achievement of students in classroom r in schools s ($\mathbf{y}_{rs}$) is given by:

$$\mathbf{y}_{rs} = \alpha + \gamma \mathbf{J}_{rs}\mathbf{a}_{rs} + \beta \mathbf{X}_{rs} + \lambda_s + \varepsilon_{rs}. \tag{13}$$

where $\mathbf{J}_{rs} = (n_{rs}-1)^{-1}(\iota_{n_{rs}}\iota'_{n_{rs}} - I_{n_{rs}})$ is a school-level block-diagonal matrix containing in each block the classroom-level leave-one-out mean operator, $\iota_{n_{rs}}$ is a vector of ones of size n_{rs}, $\mathbf{a}_{rs}$ is a dummy variable indicating whether the students' parents were investigated by the child protection services for abuse and neglect, and $\mathbf{X}_{rs}$ is a matrix of exogenous and predetermined controls including gender, whether family receive social assistance, weight at birth, mother's age, dwelling type, dwelling ownership and number of older siblings. λ_s collects school-level fixed-effects. The parameter of interest is γ, which links the proportion of disruptive classmates with own scholastic achievement.[25]

Table II.1 shows that peer abuse and neglect is associated with a significant and substantial effect on own verbal and numeric components of the intelligence test (henceforth, verbal and numeric ability) in sixth grade and on GPA in ninth grade (see the estimations *including* the A&N students and controlling for own A&N status in Table II.2). In general, the inclusion of one additional A&N classmate to a classroom of 20 students decreases verbal and numeric ability in sixth grades by 0.19 (=0.05*3.887) and

[25] In this conventional empirical model of social effects, γ is the reparametrization of the composite parameter $(I_r - \beta_{\bar{y}}\mathbf{J}_r)^{-1}\beta_{\bar{x}}$ in the reduced form LIMM.

0.24 (=0.05*4.881) points respectively. The estimates amount to negative changes of 3.2 percent and 3.0 percent of a SD in their respective scales. The average detrimental effect on cognitive achievement is larger among girls than among males. Our estimates imply that adding one more A&N peer to a classroom of 20 students decreases girls' verbal and numeric ability scores by 3.9 percent and 4.3 percent of a SD respectively. The fourth column in Table II.1 indicates that these effects on girls' cognitive ability linger enough to affect their school grades three years later by 5 percent of a SD.

Our results are very much in line with the results of Carrell and Hoekstra (2010) who use a similar reduced-form empirical strategy to explore whether children exposed to domestic violence exert negative externalities on their classmates' math and reading test scores for third to fifth graders in one county in Florida. Our estimates, ranging between 3 percent and 4 percent of a SD, are in the same order of magnitude as what they document (i.e., 2.5 percent of a SD).

When examining the heterogeneity of these results by gender of the peer exposed to abuse and neglect, we find that while A&N males affect their male and female classmates, A&N females have negative externalities on other females but not on males. For instance, the inclusion of an additional A&N *male* student to a classroom of 20 students decreases other males' numeric ability by 6.9 percent of a SD and females' numeric ability by 3.9 percent of a SD. This contrasts with the effect of the inclusion of an additional A&N *female* student to a similar classroom, which will have no significant effect on males' numeric ability, but would have a negative effect of 4.8 percent of a SD on other females' numeric ability.

Table II.1: Effects of Abused and Neglected Peers on Cognitive Scores at Ages 13 and 16

	IQ components at 13			*Grades at 16*
	Crystallized intelligence		Fluid intelligence	
	Verbal	Numeric	Spatial	GPA in grade 9
Share of A&N peers in class				
Overall effect	-3.887	-4.881	-1.275	-43.600
	(1.976)	(2.275)	(2.383)	(20.496)
Observations	7,549	7,549	7,549	7,182
Share of A&N peers in class				
Effect on males	-2.262	-3.848	0.132	-12.936
	(2.328)	(2.539)	(2.783)	(23.889)
Effect on females	-5.432	-6.217	-2.995	-70.876
	(2.253)	(2.813)	(2.560)	(23.370)
Observations	7,549	7,549	7,549	7,182
Share of male A&N peers in class				
Effect on males	-6.039	-10.515	-0.761	-39.375
	(3.722)	(4.625)	(4.608)	(43.817)
Effect on females	-6.893	-5.901	-4.417	-77.117
	(3.110)	(4.342)	(3.801)	(37.649)
Share of female A&N peers in class				
Effect on males	1.047	1.269	1.427	15.112
	(2.993)	(3.893)	(3.742)	(33.386)
Effect on females	-5.082	-7.227	-2.563	-80.898
	(3.501)	(3.891)	(3.721)	(32.389)
Observations	7,549	7,549	7,549	7,182

Note: Data from Stockholm Birth Cohort Study. Sample excludes classes with less than seven students, schools with only one class in grade 6, schools with special education classrooms and classrooms in which no one participated in the sociometric survey. Sample restricted to individuals who were not abused and neglected themselves (Angrist, 2014). Key explanatory variables are the fraction of students in the classroom whose parents underwent an investigation for abuse and neglect by the child protection services (CWC) and its interaction with a dummy for student i being female. Apart from the dummy for being female, all regressions include school fixed effects, whether family receive social assistance, weight at birth, mother's age, dwelling type, dwelling ownership and number of older siblings. Standard errors, shown in parentheses, are clustered at the school level.

Table II.2: Effects of Own A&N and A&N Peers on Cognitive Scores at Ages 13 and 16

	IQ components at 13			*Grades at 16*
	Crystallized intelligence		Fluid intelligence	
	Verbal	Numeric	Spatial	GPA in grade 9
Share of A&N peers in class				
Own A&N	-1.156	-1.295	-0-595	-19.538
	(0.249)	(0.414)	(0.317)	(4.047)
Share of A&N	-3.870	-4.442	-1.929	-37.795
	(1.915)	(2.164)	(2.346)	(18.549)
Observations	7,995	7,995	7,995	7,592
Share of A&N peers in class				
Own A&N	-1.146	-1.291	-0.584	-19.439
	(0.250)	(0.408)	(0.305)	(4.026)
Share of A&N×Male	-2.242	-3.764	-0.020	-9.050
	(2.255)	(2.438)	(2.684)	(23.082)
Share of A&N×Female	-5.396	-5.385	-4.024	-62.670
	(2.157)	(2.619)	(2.577)	(20.877)
Observations	7,995	7,995	7,995	7,592
Share of male A&N peers in class				
Own A&N	-1.153	-1.299	-0.582	-19.449
	(0.248)	(0.402)	(0.301)	(4.060)
Share of A&N×Male	-6.942	-10.510	-0.739	-37.491
	(3.488)	(4.340)	(4.242)	(40.275)
Share of A&N×Female	-7.617	-6.023	-5.208	-66.364
	(2.985)	(4.144)	(3.683)	(34.723)
Share of female A&N peers in class				
Own A&N	-1.106	-1.241	-0.572	-19.181
	(0.248)	(0.406)	(0.305)	(4.016)
Share of A&N×Male	1.946	1.505	1.318	20.631
	(2.870)	(3.634)	(3.760)	(32.770)
Share of A&N×Female	-4.301	-5.235	-4.098	-73.748
	(3.357)	(3.642)	(3.886)	(27.804)
Observations	7,995	7,995	7,995	7,592

Note: Data from Stockholm Birth Cohort Study. Sample excludes classes with less than seven students, schools with only one class in grade 6, schools with special education classrooms and classrooms in which no one participated in the sociometric survey. Note that the sample includes the individuals who were not abused and neglected themselves. Key explanatory variables are a dummy for own exposure to abuse and neglected, the fraction of students in the classroom whose parents underwent an investigation for abuse and neglect by the child protection services (CWC) and its interaction with a dummy for student i being female. Apart from the dummy for being female, all regressions include school fixed effects, whether family receive social assistance, weight at birth, mother's age, dwelling type, dwelling ownership and number of older siblings. Standard errors, shown in parentheses, are clustered at the school level.

III Structural estimation results with gender-specific adjacency matrices

In this appendix, we present the results of the structural LIMM using gender-specific adjacency matrices. They provide insights into how disruptive peer effects propagate throughout the classroom's networks differently, depending on the network members' gender. The top panel of Table III.1 reports the results for male networks and the bottom one for female networks. Matrix $\mathbf{G}^S\mathbf{y}$ represents the adjacency matrix for gender S from an otherwise similar structural LIMM as in equation (2).

We find evidence suggesting that the *endogenous effect* is an important driver of disruptive peers' effect on males but not on females. Furthermore, the *contextual effect* of males seems to be a substantial force in deterring numeric skill acquisition for both genders (the point estimates on females are substantially larger but less precise). These results provide interesting insights on how disruptiveness affects peers' numeric and verbal development differently by gender.

Table III.1: Structural Estimation Results With Gender-Specific Adjacency Matrices

	IQ Components at 13						*Grades at 16*	
	Verbal		Numeric		Spatial		GPA in grade 9	
	2SLS	2SLS+CF	2SLS	2SLS+CF	2SLS	2SLS+CF	2SLS	2SLS+CF
Males								
$\mathbf{G}^M\mathbf{y}$	0.824	0.853	0.538	0.585	1.832	1.199	0.333	0.384
	(0.281)	(0.341)	(0.225)	(0.338)	(1.422)	(1.113)	(0.287)	(0.404)
$\mathbf{G}^F\mathbf{y}$	0.094	0.118	0.171	0.164	0.277	0.228	0.044	0.019
	(0.110)	(0.140)	(0.141)	(0.181)	(0.281)	(0.235)	(0.070)	(0.112)
$\mathbf{a}$	-0.133	-0.263	-0.281	-0.253	-0.024	-0.507	-27.048	-26.821
	(0.542)	(0.551)	(0.751)	(0.790)	(0.893)	(0.656)	(6.739)	(7.056)
$\mathbf{G}^M\mathbf{a}$	-0.758	-1.075	-3.280	-3.365	0.751	-0.154	-20.654	-21.786
	(1.097)	(0.943)	(1.149)	(1.209)	(2.489)	(1.239)	(14.686)	(16.263)
$\mathbf{G}^F\mathbf{a}$	1.883	1.306	-1.376	-1.276	-2.006	-1.253	40.788	49.917
	(4.055)	(3.743)	(4.956)	(4.645)	(6.765)	(5.201)	(35.283)	(46.734)
Observations	3,873	3,873	3,882	3,882	3,873	3,873	3,828	3,828
Females								
$\mathbf{G}^M\mathbf{y}$	0.171	0.190	0.398	0.412	0.885	0.706	-0.157	-0.190
	(0.260)	(0.342)	(0.396)	(0.631)	(0.908)	(0.508)	(0.139)	(0.209)
$\mathbf{G}^F\mathbf{y}$	-0.308	-0.407	-0.300	-0.290	-3.170	-1.773	0.129	0.026
	(1.109)	(1.322)	(1.280)	(1.562)	(3.751)	(1.758)	(0.749)	(0.919)
$\mathbf{a}$	-2.410	-2.450	-2.471	-2.497	-0.945	-1.052	-35.309	-36.816
	(0.450)	(0.555)	(0.525)	(0.580)	(1.077)	(0.814)	(8.971)	(9.765)
$\mathbf{G}^M\mathbf{a}$	-2.059	-2.424	-7.064	-7.169	-11.913	-9.717	60.986	60.933
	(5.710)	(7.790)	(6.400)	(10.424)	(15.635)	(8.246)	(30.725)	(43.235)
$\mathbf{G}^F\mathbf{a}$	-2.443	-2.776	-1.919	-1.901	-2.670	-1.294	-29.408	-35.838
	(2.946)	(3.239)	(3.433)	(3.957)	(4.077)	(1.912)	(37.643)	(44.706)
Observations	4,091	4,091	4,092	4,092	4,091	4,091	4,050	4,050

Note: Data from Stockholm Birth Cohort Study. Table present the estimates of structural model (2) allowing for gender-epecific adjacency matrices. The coefficients associated with $\mathbf{G}^M\mathbf{y}$ and $\mathbf{G}^F\mathbf{y}$ represent the endogenous peer-effect caused by male students and female students, respectively. The coefficients associated with $\mathbf{G}^M\mathbf{a}$ and $\mathbf{G}^F\mathbf{a}$ represent the exogenous peer-effects caused by male and female abused and neglected classmates, respectively. Estimations include school fixed effects. Sample excludes classes with less than seven students, schools with only one class in grade 6, schools with special education classrooms and classrooms in which no one participated in the sociometric survey. 2SLS models use clustered standard errors at the classroom level. 2SLS+CF models use clustered-at-the-classroom-level bootstrapped standard errors.

IV Monte Carlo simulation of networks to aid interpretation of structural results

In this Section, we explain in detail the simulations presented in Figures 1 and 2 in the main text that are based on the structural results of the structural LIMM of equation (2). The goal of these Monte Carlo simulations is to aid the interpretation of the estimates reported in Tables 4 and III.1. The simulations provide a way to summarize the estimates reported there in a way that incorporates all possible feedback loops that stem from interconnections within social clusters and cliques. The intuition for the simulation is the following: we randomly create a network $\mathcal{C}_r^{(1)}$, assign A&N peers within that network (i.e., $a \in \mathbf{a}_r^{(1)}$ s.t. $a = 1$) and, based on the estimated parameters in Tables 4 or III.1 (i.e., $\hat{\beta}$, $\hat{\beta}_{\bar{y}}$ and $\hat{\beta}_{\bar{x}}$) and the shape and count of the network's edges ($\mathbf{G}_r^{(1)}$), let the model—and therefore, the location relative to the A&N peer—yield the expected outcome $\hat{y}^{(1)}$ for each individual.

Of course, the shape of $\mathcal{C}_r^{(1)}$ is critical in defining the expected outcome $\hat{y}^{(1)}$. For this reason, we take 1,000 random draws of social networks (i.e., $\mathcal{C}_r^{(i)}$ for $i = 1, \ldots, 1000$) that, on average, match the characteristics of the average classroom we observe in our data. Specifically, we calibrate a distribution of dyad linking probabilities p so that the simulated networks have, in expectation, the same number of edges (40.4) and clustering (0.41) as the average classroom in our data. We further impose the simulated networks to have 24 nodes (average class size), two A&N peers (average number of such peers in a classroom), and genders equally distributed.

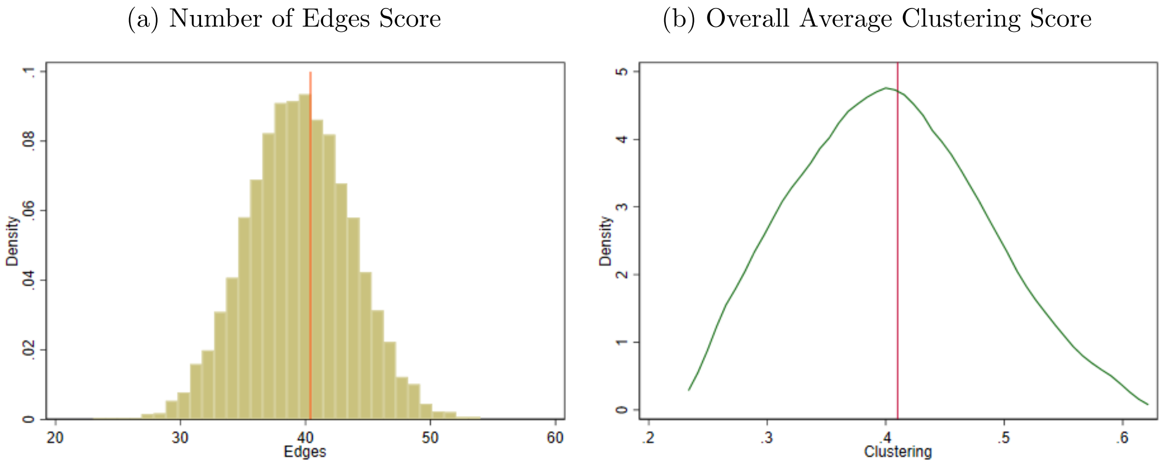
The calibration of the dyad linking probabilities uses the facts that children are more likely to befriend classmates of the same gender (only 3 percent of the friendships are cross-gender), and that even within genders, there is some clustering going on where some students have a higher probability of becoming friends than others. Thus, we define three linking probabilities p_h, p_m and p_l as indicated in matrix $\mathbf{P}$:

$$
\mathbf{P} = \begin{pmatrix}
0 & p_h & \cdots & p_h & p_m & \cdots & p_m & p_l & \cdots & p_l & p_l & \cdots & p_l \\
 & 0 & \ddots & \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\
 & & 0 & p_h & p_m & \cdots & p_m & p_l & \cdots & p_l & p_l & \cdots & p_l \\
 & & & 0 & p_h & \cdots & p_h & p_l & \cdots & p_l & p_l & \cdots & p_l \\
 & & & & 0 & \ddots & \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\
 & & & & & 0 & p_h & p_l & \cdots & p_l & p_l & \cdots & p_l \\
 & & & & & & 0 & p_h & \cdots & p_h & p_m & \cdots & p_m \\
 & & & & & & & 0 & \ddots & \vdots & \vdots & \dots & \vdots \\
 & & & & & & & & 0 & p_h & p_m & \cdots & p_m \\
 & & & & & & & & & 0 & p_h & \cdots & p_h \\
 & & & & & & & & & & 0 & \ddots & \vdots \\
 & & & & & & & & & & & 0 & p_h \\
 & & & & & & & & & & & & 0
\end{pmatrix}
$$

where nodes are sorted by gender. Without loss of generality, let the first half of nodes be males and the second half be females. Then, p_l is the probability two nodes of different gender build a friendship link. p_h and p_m are the probabilities of befriending a student from the same gender. We assume that $p_h > p_m$ as it represents the fact that even within gender, students friendships tend to cluster. Therefore, $p_h > p_m > p_l$ as cross-gender friendships are rare events.

In order to match the network characterictics of the average classroom, we chose $p_h = 0.335$, $p_m = 0.03$ and $p_l = 0.005$. Figures IV.1 show that the networks we simulate based on these probabilities produce distributions of networks characteristics centered at the target measures of number of edges and overall average clustering.

Figure IV.1: Simulated Networks Characteristics



Note: Distributions of characteristics of the simulated networks. The vertical red lines represent the target characteristics of the average network we observe in the data. Bottom and top 1 percent in the individual clustering distribution were dropped.

Based on these probabilities, we randomly generate 1,000 networks $\mathcal{C}_r^{(1)} \dots \mathcal{C}_r^{(1000)}$. Then, we randomly allocate the A&N peers within those networks $\mathbf{a}_r^{(1)} \dots \mathbf{a}_r^{(1000)}$, and generate the corresponding outcomes $\hat{\mathbf{y}}_r^{(1)} \dots \hat{\mathbf{y}}_r^{(1000)}$ based on model (2) and the estimates in Tables 4 and III.1. Finally, we summarize the results of our simulations by averaging the outcomes by the geodesic to the closest A&N student (i.e., $\mathbb{E}[\hat{\mathbf{y}}_r^{(i)} | \, ||\mathbf{r}|| = \rho]$ for $\rho \in \{1, 2, 3, 4, 5\}$).

V Robustness

V.1 Stability of the coefficients in the distance equation.

Table V.1: Testing Stability of the Coefficients in the Model on Effects of Distance to A&N Peers on Cognitive Scores

	IQ Components at 13			*Grades at 16*
	Verbal	Numeric	Spatial	GPA in grade 9
Panel A				
A&N in class	-1.522	-1.422	-1.084	-0.187
	(0.629)	(0.657)	(0.763)	(0.108)
Not connected	1.427	1.320	1.264	0.168
	(0.581)	(0.684)	(0.769)	(0.108)
Distance to A&N	0.491	0.524	.477	0.076
	(0.221)	(0.265)	(0.289)	(0.042)
Fract. classmates with same Z	YES	YES	YES	YES
Panel B				
A&N peer in class	-1.444	-1.397	-1.039	-0.169
	(0.612)	(0.645)	(0.771)	(0.110)
Not connected	1.361	1.288	1.244	0.147
	(0.561)	(0.675)	(0.777)	(0.107)
Distance to A&N peer	0.465	0.508	0.446	0.067
	(0.216)	(0.262)	(0.291)	(0.043)
Fract. classmates with same Z	YES	YES	YES	YES
Observations	7,464	7,464	7,464	7,101

Note: Note: Data from Stockholm Birth Cohort Study. Sample excludes classes with less than seven students, schools with only one class in grade 6, schools with special education classrooms and classrooms in which no one participated in the sociometric survey. *A&N in class* takes on value one if there is a student in the classroom whose parents underwent an investigation for abuse and neglect by the child protection services (CWC) and zero otherwise; *Not connected* takes on value one if despite there being an abused and neglected peer in the classroom he/she does not belong to the social network of the student; and *Distance to A&N* stands for the path length between the student and the closest abused and neglected peer. All regressions include school fixed effects, a dummy for being female, whether family receive social assistance, birth weight, mother's age, dwelling type, dwelling ownership and number of older siblings and the averages of classmates (Panel A) and average of friends (Panel B) who are similar to student i in terms of each particular homophily variables of vector Z_i of the link formation estimating equation (9), reported in Table I.2. Further, the estimated degree heterogeneity $\hat{\theta}_i$ for student i is treated as an included instrument and is hence included in both the first stages and the second stage. Standard errors, shown in parentheses, are clustered at the school level.

V.2 Truncation in the nomination of friends

Table V.2: Effects of Distance to A&N Peers on Cognitive Scores at Ages 13 and 16. Full Sample Adding Friendship Nominations

	IQ Components at 13						Grades at 16	
	Verbal		Numeric		Spatial		Marks 9	
	Top-Coded	LIML	Top-Coded	LIML	Top-Coded	LIML	Top-Coded	LIML
Adding the closest friend that is not a friend								
A&N in class	-1.167	-1.799	-0.953	-1.542	-0.677	-1.148	-0.147	-0.237
	(0.416)	(0.664)	(0.522)	(0.693)	(0.466)	(0.809)	(0.071)	(0.110)
Not connected		1.488		1.318		1.142		0.212
		(0.598)		(0.708)		(0.816)		(0.113)
Distance to A&N	0.284	0.753	0.255	0.689	0.237	0.608	0.043	0.113
	(0.118)	(0.281)	(0.148)	(0.336)	(0.132)	(0.365)	(0.020)	(0.051)
Observations	7,464	7,464	7,464	7,464	7,464	7,464	7,101	7,101
Adding a friend's friend								
A&N in Class	-1.265	-1.727	-1.032	-1.542	-0.746	-1.235	-0.161	-0.261
	(0.447)	(0.654)	(0.560)	(0.659)	(0.500)	(0.795)	(0.076)	(0.112)
Not connected		1.643		1.349		1.380		0.239
		(0.583)		(0.672)		(0.800)		(0.112)
Distance to A&N	0.296	0.644	0.263	0.640	0.243	0.594	0.045	0.116
	(0.121)	(0.258)	(0.151)	(0.300)	(0.135)	(0.336)	(0.021)	(0.049)
Observations	7,464	7,464	7,464	7,464	7,464	7,464	7,101	7,101
Adding a neighbor as friend								
A&N in class	-1.471	-1.561	-1.198	-1.260	-0.890	-1.226	-0.192	-0.220
	(0.515)	(0.643)	(0.644)	(0.678)	(0.575)	(0.794)	(0.088)	(0.107)
Not connected		1.269		1.120		1.128		0.184
		(0.577)		(0.674)		(0.846)		(0.110)
Distance to A&N	0.418	0.599	0.366	0.510	0.335	0.616	0.063	0.099
	(0.167)	(0.258)	(0.209)	(0.311)	(0.187)	(0.332)	(0.029)	(0.047)
Observations	7,464	7,464	7,464	7,464	7,464	7,464	7,101	7,101
Adding a friend if popular								
A&N in Class	-1.224	-1.771	-1.001	-1.638	-0.720	-1.125	-0.156	-0.233
	(0.435)	(0.645)	(0.546)	(0.678)	(0.487)	(0.783)	(0.074)	(0.109)
Not connected		1.590		1.376		1.258		0.213
		(0.585)		(0.689)		(0.798)		(0.110)
Distance to A&N	0.294	0.679	0.263	0.694	0.243	0.544	0.045	0.103
	(0.121)	(0.256)	(0.152)	(0.309)	(0.135)	(0.329)	(0.021)	(0.047)
Observations	7,464	7,464	7,464	7,464	7,464	7,464	7,101	7,101

Note: Data from SBC. Sample excludes classes with less than seven students, schools with only one class in grade 6, schools with special education classrooms and classrooms in which no one participated in the sociometric survey. Sample restricted to individuals who were not abused and neglected. Table present the estimates of the distance model with additional friendship nominations for students who nominated three friends (i.e., the maximum number). In the top panel labeled *Adding the closest friend that is not a friend*, we include an additional friendship nomination to every student. In the next panel labeled *Adding a friend's friend*, the added friendship is a friend's friend. In panel labeled *Adding a neighbor* the additional friendship nomination is a student who lives in the same block. In the bottom panel labeled *Adding a friend if popular* we include an additional friendship nomination to every student who nominated three friends *and* has an estimated degree heterogeneity above the 50th percentile. In all cases, the additional friendship nomination corresponds to the classmate with the highest predicted probability of friendship among those who are not reported as friends. *A&N in class* takes on value one if there is a student in the classroom with abusive/neglectful parents according to the CWC and zero otherwise; *Not connected* takes on value one if the potential A&N peer in the classroom does not belong to the same network component of the student; and *Distance to A&N* stands for the path length between the student and the closest A&N peer. All regressions include school fixed effects and a dummy for being female. Further, the estimated degree heterogeneity $\hat{\theta}_i$ for student i is treated as an included instrument and is hence included in both the first stages and the second stage. Standard errors, shown in parentheses, are clustered at the school level.

Table V.3: Structural Estimation Results: Adding Friendship Nominations

	IQ Components at 13							
	Verbal				Numeric			
	(1) All	(2) Fr of Fr	(3) Neighbor	(4) Top 50	(1) All	(2) Fr of Fr	(3) Neighbor	(4) Top 50
Gy	0.624	0.649	0.720	0.575	0.351	0.473	0.474	0.254
	(0.347)	(0.299)	(0.336)	(0.305)	(0.439)	(0.322)	(0.363)	(0.416)
a	-1.267	-1.287	-1.264	-1.313	-1.667	-1.529	-1.588	-1.744
	(0.399)	(0.363)	(0.374)	(0.372)	(0.556)	(0.485)	(0.480)	(0.547)
Ga	-0.486	-0.270	-0.195	-0.650	-2.229	-1.502	-1.946	-2.474
	(1.005)	(0.927)	(0.986)	(0.915)	(1.314)	(1.227)	(1.139)	(1.350)
Observations	7,964	7,964	7,964	7,964	7,974	7,974	7,974	7,974

Note: Data from Stockholm Birth Cohort Study. Table present the estimates of structural model (2) with the inclusion of friendship nominations additional to the ones reported in the data for students who nominated three friends (i.e., the maximum). In Columns labeled *All*, we include an additional friendship nomination to every student. In Columns labeled *Fr of Fr*, we include an additional friendship where the friendship nomination given is a friend's friend. In Columns labeled *Neighbor* the additional friendship nomination is a student who lives in the same block. In Columns labeled *Top 50* we include an additional friendship nomination to every student who nominated three friends *and* has an estimated degree heterogeneity above the 50th percentile. In all cases the additional friendship nomination corresponds to the classmate with the highest predicted probability of friendship among those who are not reported as friends. The coefficients associated with **Gy** represent the endogenous peer-effect. The coefficients associated with **Ga** represent the exogenous peer-effects. We use $\mathbf{G^2a}$ and $\mathbf{G^3a}$ as instruments for **Gy**, plus a control function approach described in Johnsson and Moon (2021). Sample excludes classes with less than seven students, schools with only one class in grade 6, schools with special education classrooms and classrooms in which no one participated in the sociometric survey. Estimations include school fixed-effects. Robust standard errors in parentheses.

V.3 Placebo regressions

Table V.4: Placebo test of the Model on Effects of Distance to A&N Peers on Cognitive Scores

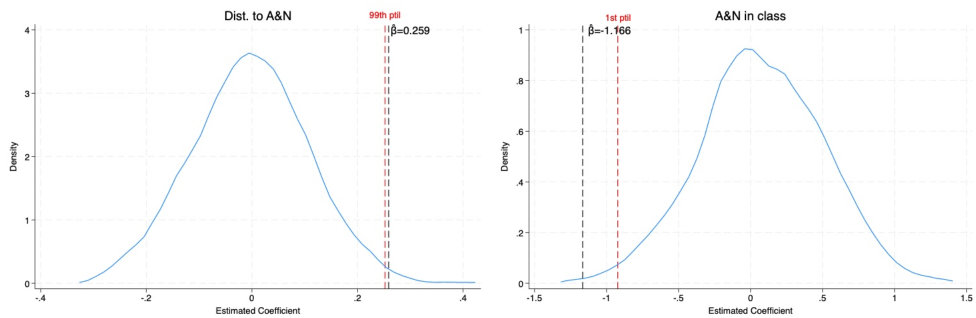
	Duration of mother's postpartum stay	Height at birth	Number of rooms in the apartment (1960)	Number of individuals in household (1960)	High SES father	Height measured at enlistment
A&N in class	0.171	-0.030	-0.020	0.019	-0.006	-0.443
	(0.325)	(0.259)	(0.099)	(0.128)	(0.025)	(0.938)
Not connected	0.157	0.119	0.052	0.005	0.015	0.515
	(0.328)	(0.143)	(0.098)	(0.123)	(0.027)	(0.985)
Dist. to A&N	0.040	-0.020	0.001	-0.014	0.004	0.183
	(0.134)	(0.057)	(0.038)	(0.046)	(0.010)	(0.369)
Observations	6,025	6,018	7,042	7,008	7,464	3,269

Note: Data from Stockholm Birth Cohort Study. Sample excludes classes with less than seven students, schools with only one class in grade 6, schools with special education classrooms and classrooms in which no one participated in the sociometric survey. *A&N in class* takes on value one if there is a student in the classroom whose parents underwent an investigation for abuse and neglect by the child protection services (CWC) in the classroom and zero otherwise; *Not Connected* takes on value one if despite there being a disruptive peer in the classroom he/she does not belong to the social network of the student; and *Dist. to A&N* stands for the path length between the student and the closest abused and neglected peer. All regressions include school fixed effects, a dummy for being female, whether family receive social assistance, birth weight, mother's age, dwelling type, dwelling ownership and number of older siblings. Standard errors, shown in parentheses, are clustered at the school level.

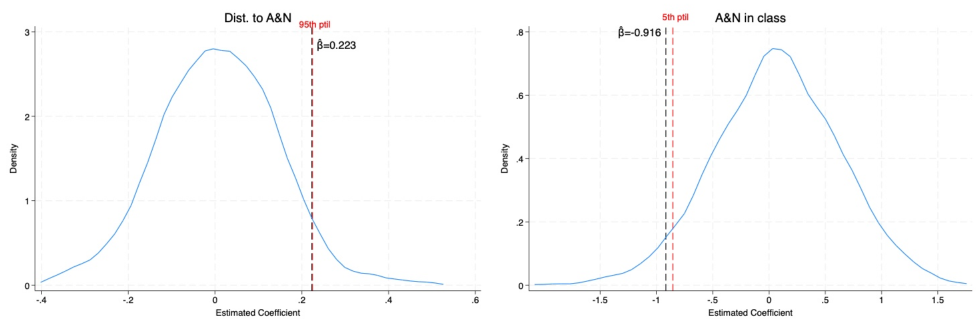
V.4 Placebo treatment analyses

Figure V.1: Distribution of the Coefficients of Placebo Treatment Analyses

(a) Verbal



(b) Numeric



Note: The figures plot the kernel distributions of coefficients from 1,000 estimations where treatment (i.e., being a A&N peer) is randomly assigned. The estimated models are those in which the distance between two not-connected peers is top-coded to be equal to the network diameter plus one. The black dashed lines indicate the corresponding coefficients from our actual estimations presented in Table 2 in our main manuscript. The red dashed lines indicate extreme quantiles for reference. Data from Stockholm Birth Cohort Study. Sample excludes classes with less than seven students, schools with only one class in grade 6, schools with special education classrooms and classrooms in which no one participated in the sociometric survey. Sample restricted to individuals who were not abused and neglected themselves. All regressions include school fixed effects, a dummy for being female, whether family receive social assistance, birth weight, mother's age, dwelling type, dwelling ownership, number of older siblings, and degree heterogeneity $\hat{\theta}_i$.

VI Structural results: First stage

Table VI.1: Structural Estimation: First Stage Results

	Gy *of IQ Components at 13*						**Gy** *of Grades at 16*	
	Verbal		Numeric		Spatial		GPA in grade 9	
$\mathbf{X}$	-0.272	-0.216	-0.309	-0.280	-0.038	-0.014	-2.697	-2.646
	(0.123)	(0.123)	(0.148)	(0.147)	(0.138)	(0.138)	(1.789)	(1.784)
$\hat{\mathbf{G}}(\mathbf{Z})\mathbf{X}$	-3.562	-17.357	-6.590	-13.745	-3.288	-9.429	-56.412	-68.038
	(2.422)	(0.946)	(2.909)	(1.133)	(2.715)	(1.058)	(35.141)	(13.706)
$\hat{\mathbf{G}}(\mathbf{Z})^2\mathbf{X}$	-39.324		-20.386		-17.507		-32.949	
	(6.360)		(7.634)		(7.128)		(91.706)	
$\hat{\mathbf{G}}(\mathbf{Z})^3\mathbf{X}$	57.855	26.748	34.299	18.176	28.434	14.586	45.954	19.976
	(5.333)	(1.774)	(6.400)	(2.125)	(5.977)	(1.984)	(76.717)	(25.639)
Observations	7,481	7,481	7,493	7,493	7,481	7,481	7,153	7,153
F	97.31	116.4	45.50	58.24	21.79	27.03	22.43	29.87
Pr $> F$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Note: Data from Stockholm Birth Cohort Study. Table present the first stage estimates of structural results shown in Table 4. The coefficients associated with $\mathbf{GX}$ represent the exogenous peer-effects. $\hat{\mathbf{G}}(\mathbf{Z})$ is the predicted adjacency matrix obtained from (9). Sample excludes classes with less than seven students, schools with only one class in grade 6, schools with special education classrooms and classrooms in which no one participated in the sociometric survey. Estimations include school fixed-effects. Constant not shown. Robust standard errors in parentheses.

VII Attrition simulations

As explained in Section 2, one issue with our data is that we do not observe the nominations made to classmates in sixth grade who were not born in 1953 and hence did not belong to the studied birth cohort. They comprise about 9 percent of the students in sixth grade. Although this type of attrition is relatively small in our case (Boucher and Houndetoungan, 2023), it can potentially pose some challenges as the retained students are more likely to be A&N. Our data (1953 cohort) shows that they are 7 percentage points more likely to be A&N than cohort members in the right grade for age. Thus, assuming that the retainers born in 1952 are similar to the observed retainers born in 1953 (being in fifth grade in 1966, and hence missing the sixth grade survey), we risk missing some classmates who are more likely to be A&N children.

In this section, we perform simulation exercises to infer how would the above-referenced attrition would affect our results. In each simulation, we do the following routine:

1. From our data, we take the observed A&N children of the 1953 cohort that are not in the right (sixth) grade for their age in 1966 and let them proxy the older (cohort 1952) retainers that are actually in sixth grade in 1966. We randomly assign these students to our observed sixth grade classrooms, following the observed distribution of A&N incidence at the school level.

2. Within their assigned classroom, we assign these students to be friends with their most likely 'classmates'. Those are the ones with the highest predicted probabilities of friendship based on their characteristics and according to the friendship formation model.

3. We calculate new adjacency matrices including these new synthetic friendships, and calculate new geodesics.

4. We run our empirical models on these synthetic data.

We repeat this computational intensive exercise 150 times. We expect this procedure to yield a lower bound for our estimates. That is because we are actually assigning A&N peers where they are not likely to be, and thus expecting to see an effect where there is no treatment. Furthermore, we are effectively shortening the geodesic between students and the A&N peer, and decreasing the number of classrooms without A&N students. For instance, while 34.5 percent if the classrooms in our data have no A&N peers, in the simulated data, on average, only 30.9 percent of the classrooms lack an A&N student.

The results of our simulation exercises in Table VII.1 show that, as expected, the estimates with the synthetic data are a lower bound. However, importantly, the negative relation between the size of the peer effect and distance to the A&N peer persists.

Table VII.1: Effects of Distance to Abused and Neglected Peers Simulating Classrooms with Students From Other Cohorts

	IQ Components at 13			*Grades at 16*
	Verbal	Numeric	Spatial	GPA grade 9
A&N in class	-1.025	-1.472	-0.598	-0.091
	(0.634)	(0.697)	(0.685)	(0.087)
Not connected	0.572	1.057	0.463	0.034
	(0.538)	(0.630)	(0.631)	(0.081)
Dist to A&N	0.312	0.473	0.267	0.028
	(0.216)	(0.244)	(0.250)	(0.032)

Note: Data from Stockholm Birth Cohort Study. Sample extended by including A&N children of the 1953 cohort that are not in the right grade for their age. They are randomly assigned to classrooms and friends within those classrooms. Sample excludes classes with less than seven students, schools with only one class in grade 6, schools with special education classrooms and classrooms in which no one participated in the sociometric survey. Sample restricted to individuals who were not abused and neglected themselves. All regressions include school fixed effects, a dummy for being female, whether family receive social assistance, birth weight, mother's age, dwelling type, dwelling ownership and number of older siblings. Further, the estimated degree heterogeneity $\hat{\theta}_i$ for student i is treated as an included instrument. $\hat{\theta}_i$ for the A&N children simulated into classrooms are centered at the population mean. Reported estimates are the result of 150 repetitions of the LIML estimation of the distance model outlined in equation (8) in the main text. Standard errors, shown in parentheses, are clustered at the school level.