Effects of Disruptive Peers in Endogenous Social Networks*

Torsten Santavirta† Miguel Sarzosa‡

June 12, 2019

Abstract

This study explores the role of social networks in shaping the diffusion (i.e., spread and heterogenous impact) of the negative classroom externalities generated by disruptive peers. We recognize that social networks are endogenous and model friendship formation based on homophily. Using the Stockholm Birth Cohort 1953, we find that classmates exposed to abuse and neglect have negative effects on peers' cognitive achievement at ages 13, 16 and 19. We find stronger effects for those students who are socially closer to the disruptive peers and a fade out that limits the reach of the externality to a path length of three edges. Finally, social networks provide instruments that allow us to estimate the structural parameters of the modified linear-in-means model. We find that most of the effect disruptive peers exert on their classmates is accounted for by the direct effect of their disruptiveness, leaving no scope for social multiplier effects.

JEL Classification: I21, J13, J24

*This paper was supported by the Handelsbanken Research Foundation grant number P2016-0093. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Handelsbanken Research Foundation. We would like to thank Ylva Almquist and Jannike Kjellström at Centre for Health Equity Studies (CHESS) for their help with the data construction. We are grateful to thank Eric Auerbach, Manuel Bagues, Tim Bond, Vincent Boucher, Jillian Carr, Eric Chyn, Gordon Dahl, Magne Mogstad, Kevin Mumford, Matthew Lindquist, Áureo de Paula, Michele Pellizzari, Victoria Prowse, Eirini Tatsi and the participants in the Purdue Summer Seminar Series (West Lafayette, 2017), SOFI labor seminar (Stockholm, 2018), HECER lunch seminar (Helsinki, 2018), SOLE (Toronto, 2018), University of York (York, UK, 2018), Purdue HDFS Colloquium (West Lafayette, 2018), Ohio State University (Columbus, 2018), LACEA Meetings (Guayaquil, Ecuador, 2018), 3rd Applied Micro Work Shop (Bolzano, Italy, 2019), 5^{th} Conference on Network Science and Economics (IU, Bloomington, 2019) for useful comments.

†Institute for Housing and Urban Research, Uppsala University and Swedish Institute for Social Research, Stockholm University. Box 514, 751 20 Uppsala, 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

Existing studies document causal negative externalities of disruptive peers on achievement scores in elementary school (Aizer, 2008; Carrell and Hoekstra, 2010), and show that the consequences persist into the early career labor market, at least until age 26 (Carrell et al., 2018). Despite these seminal contributions, we know very little about these important externalities beyond their existence and total magnitude. In particular, studies that uncover the complex mechanisms that govern the spread and heterogeneity of peer effects in educational settings are scarce (Jackson and Yariv, 2011; Sacerdote, 2011; Carrell et al., 2013; de Paula, 2017).[1] There is a considerable upside to this kind of research as it could guide policy makers towards interventions that dissipate and contain the spread of the externality (Graham, 2018; Manski, 2013).

This paper exploits data on friendship nominations within the classroom to explore the role of social networks in shaping the diffusion (i.e., spread and heterogenous impact) of the negative externalities generated by disruptive peers.[2] We do so in two ways. First, we derive and estimate a model in which social distance between the disruptive peer and each student affects the size of the externality. Second, building on the strategy of Bramoullé et al. (2009) and De Giorgi et al. (2010), we identify the structural parameters of the linear-in-means model using the exogeneity of friends' characteristics and the existence of intransitive triads of friends (i.e., that for some individuals, the friends

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

[2]In this regard, our paper relates to the literature that examines the role of networks in shaping the diffusion of information and behaviors (Calvó-Armengol et al., 2009; Jackson and Yariv, 2011; Banerjee et al., 2013; Schennach, 2018). In particular, our paper extends a sub-branch of the literature that focuses on the role networks have in the diffusion of the effects of treatments and interventions in education settings. List et al. (2018) find that spillovers of an early childhood education intervention followed the children's social networks. Opper (2019) shows that middle school students benefit from their friends (proxied by children with the same gender and race) having had good elementary school teachers.

of friends are not own friends). This allows us to disentangle the effect that emanates from peers' outcomes (endogenous) from the direct effect of the peers' disruptiveness (exogenous) under certain assumptions (Manski, 1993). Throughout the analysis we recognize that networks in classrooms result from students' active choices and are therefore endogenous. To overcome this confounding problem, we model the link formation process based on homophily (i.e., the fact that people tend to befriend others that are similar to themselves) (Goldsmith-Pinkham and Imbens, 2013; Hsieh and Lee, 2015). We do so in terms of predetermined characteristics of the student such as gender, the neighborhood block the parents lived in when the student was born, parental SES at the time of birth, and prenatal and perinatal characteristics.

Following the existing literature and in the interest of exogeneity, we proxy child's disruptiveness in school by exposure to parental abuse and neglect during childhood (Carrell and Hoekstra, 2010; Carrell et al., 2018).[3] We will henceforth refer to these students as "disruptive peers". Identification of disruptive classmates through their parents' behavior combined with random (within-school) allocation of these students to classrooms circumvents a number of potential confounders that would bias our peer effect estimates. In this sense, part of our identification strategy follows Hoxby (2000) in exploiting the variation in classroom composition within schools to identify the peer effects.[4] It is facilitated by the institutional framework that administered assignment

[3]Existing literature in Psychology and Economics has shown that child maltreatment significantly contributes to a number of 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, 2015). Table A.1 in Appendix A shows empirically that children exposed to abuse and neglect in our data have substantially lower cognitive outcomes, are more likely to have adjustment problems and to engage in risky behaviors relative to both regular students and other subgroups of students from disadvantaged backgrounds. This is in line with existing evidence on the effects of childhood exposure to abuse, neglect and domestic violence on life outcomes (Currie and Widom, 2010; Aizer, 2011; Doyle Jr, 2007, 2008).

[4]Hoxby (2000) introduced the idea of using cohort-to-cohort population variation in classroom composition to study peer effects in elementary school. Researchers have used variations of this approach to quantify classroom peer effects exerted by children with attention deficit disorder (ADD) (Aizer,

to Swedish schools in the 1960's. In particular, students attended the nearest school in the neighborhood and tracking based on ability or background was strongly discouraged resulting in randomly formed classrooms rosters within elementary schools, at least with respect to the share of disruptive students.

We explore a unique longitudinal dataset that follows up the entire cohort of children born in Stockholm, Sweden in 1953. The Stockholm Birth Cohort Study (SBC) links survey data with administrative records including birth records, population censuses and the universe of reports of having been investigated by the Swedish child protection services agency (henceforth CWC, for the direct English translation, Child Welfare Committee) for suspicion of child abuse and neglect. To identify students who have experienced abuse and neglect at home we use those CWC investigations that resulted in measures against the parents. Importantly, the school survey conducted in Spring term of sixth grade contains friendship nominations and tests of verbal, numeric and spatial components of intelligence.

To the best of our knowledge, this is the first study to quantify heterogeneous peer effects within reference groups defined by the endogenous social network structure that stems from student friendships.[5] The novelty here lies in using information on endogenous within-classroom networks not only as an identification tool, but as a source of variation that allows us to understand better the diffusion of peer effects.[6] Key to the

2008), by migrants (Gould et al., 2009), by children linked to domestic violence (Carrell and Hoekstra, 2010; Carrell et al., 2018), by gender composition (Black et al., 2013), by boys with female sounding names (Figlio, 2005), by low-ability students (Lavy et al., 2011), by children with college-educated mothers and students from minority groups (Bifulco et al., 2011), and the effect of skill composition on reading scores (Ammermueller and Pischke, 2009).

[5] de Paula (2017) provides an extensive discussion on studies that model the network formation and outcome equation simultaneously. Further recent innovative peer effect studies solving the correlated effects problem of endogenous network formation include Comola and Prina (2015) in the context of informal financial networks, Hsieh and Kippersluis (2018) in the context of smoking initiation and König et al. (2018) in the context of R&D networks.

[6] De Giorgi et al. (2010), Lin (2010) and Patacchini et al. (2017) use empirical strategies that combine

identification is coming to terms with endogenously formed networks. We estimate a model of endogenous link formation based on homophily that allows for degree heterogeneity. Finally, we are the first to identify the parameters of the linear-in-means model in the disruptive peer context.

In the reduced-form analysis, we document that the share of disruptive peers in the classroom in sixth grade has a negative causal effect on own cognitive test scores and grades. The size of this effect is very close to that reported by Carrell and Hoekstra (2010) for students in a county in Florida, US. In the analysis of how the externality permeates through the social network, we find that the disruptive peer exerts the largest negative effect on the cognitive outcomes of her closest friends. The externality fades out at a rate proportional to the social distance to the disruptive peer. If the shortest path length (geodesic) between the student and the disruptive peer is more than two friendship links (edges) in the social network, the statistically significant negative peer effect will have faded out completely. Finally, the structural estimation and identification of the parameters of the linear-in-means model show that disruptive peers affect their friends' outcomes directly (contextual effects) absent of a social multiplier effect that would go through peer's cognitive outcomes (endogenous effects). By documenting that the negative externality of having a disruptive peer is stronger among the students more closely related to her combined with showing evidence of the presence of contextual effects of disruptive peer characteristics (but no endogenous effects of peer's outcomes), this study brings forth recommendations for designing optimal educational policy. As in Aizer (2008), our results suggest that there is little scope for allocative policies that attempt to minimize the negative spillovers by evening out the disruptive

the use of intransitive triads with social cliques fixed-effects to separately identify the contextual and the endogenous effects in education settings. Patacchini et al. (2017) do provide some effect heterogeneity depending on the length of the friendship. They find that strong friendships—those that last for more than a year—have greater scope of influencing behaviors.

peers across classrooms. Instead, more support to classes and schools with disruptive peers may be more effective.

The remainder of this article is organized as follows. Section 2 outlines the theoretical framework. Section 3 describes the data, construction of the analytic sample and measures. Section 4 presents the empirical strategies of estimating both our reduced-form and structural model of peer effects and discusses the validity of the underlying identifying assumptions. Section 5 presents the main estimates of disruptive peer effects on students' cognitive outcomes and presents evidence of the mechanisms driving the negative effects of disruptive peers. Section 6 tests the robustness of our results. Section 7 concludes.

2 From negative externalities to peer effects in networks

This section describes our structural framework for inferring the presence, and identifying the type and size, of peer effects using network data. The framework builds on the linear-in-means model of social interactions (Manski, 1993). It uses network data on who is connected to whom to overcome the reflection problem and to identify the structural parameters of the model. We further leverage on the expression of the reduced-form action vector using a series expansion to propose a reduced-form model that exploits the network data to identify how much students' connectedness to the disruptive peer matters for their cognitive outcomes.

2.1 The reduced-form action vector and endogenous networks

We start from a linear best reply-function of the basic linear-in-means type of model of social interactions (Manski, 1993; Brock and Durlauf, 2001; Bramoullé et al., 2009; Blume et al., 2011)[7]:

$$y_{ri} = \beta_r + \beta X_{ri} + \beta_{\bar{y}} \bar{Y}_{r,s-i} + \beta_{\bar{x}} \bar{X}_{r,s-i} + \eta_{ri} \tag{1}$$

where an agent i in reference group r has a best reply (y_{ri}) that varies with the average actions of the members of her reference group $(\bar{Y}_{r,s-i})$, her own observed characteristics (X_{ri}) and the average characteristics of her reference group $(\bar{X}_{r,s-i})$, the unobserved reference group effect (β_r) and unobserved own characteristics (η_{ri}). Defining ι_{n_r} as a vector of ones of size n_r—the number of individuals in the r^{th} reference group—and $\mathbf{J}_r = (n_r - 1)^{-1}(\iota_{n_r}\iota'_{n_r} - I_{n_r})$, we can write (1) in matrix form

$$\mathbf{Y}_r = \beta_r \iota_{n_r} + \beta \mathbf{X}_r + \beta_{\bar{y}} \mathbf{J}_r \mathbf{Y}_r + \beta_{\bar{x}} \mathbf{J}_r \mathbf{X}_r + \eta_r, \tag{2}$$

where $\mathbf{Y}_r$ is a $n \times 1$ vector and $n = \sum_{r=1}^{\bar{r}} n_r$ is the total number of observations in all reference groups taken together. $\mathbf{J}_r$ is the $n \times n$ leave-one-out mean operator; a block diagonal matrix that we can interpret as a row-normalized socio-matrix for reference groups where all individuals are linked to one other. Solving for $\mathbf{Y}_r$ the equilibrium

[7] Most empirical work on the linear-in-means model does not rely on a theoretical model for its micro-foundations but treats it as a statistical model inspired by spatial econometrics (Blume et al., 2011). However, there are notable examples of game theoretic models of social interactions that yield linear best reply functions in the flavor of the linear-in-means model. For example, Patacchini and Zenou (2012) and Blume et al. (2015) model social interactions as emerging from social norms in a non-cooperative game and Calvó-Armengol et al. (2009) model social interactions through a game with strategic complementarities in own effort. See for Jackson and Zenou (2015) for a survey of models that provide rigorous micro-foundations for empirical analyses of the linear-in-means model.

action vector as a function of $\mathbf{J}$, $\mathbf{X}$, β_r and η_r alone yields the reduced-form:

$$\mathbf{Y}_r = \frac{\beta_r}{1-\beta_{\bar{y}}}\iota_r + (I_r - \beta_{\bar{y}}\mathbf{J}_r)^{-1}\beta\mathbf{X}_r + (I_r - \beta_{\bar{y}}\mathbf{J}_r)^{-1}\beta_{\bar{x}}\mathbf{J}_r\mathbf{X}_r + (I_r - \beta_{\bar{y}}\mathbf{J}_r)^{-1}\eta_r \tag{3}$$

Regression analysis of (3) will yield unbiased estimates of $(I_r - \beta_{\bar{y}}\mathbf{J}_r)^{-1}\beta_{\bar{x}}$ as long as $\mathbf{X}_r \perp \eta_r$, but it will not disentangle $\beta_{\bar{y}}$ from $\beta_{\bar{x}}$. This formalizes the *reflection problem* (Manski, 1993). It is not solved by the exogeneity of $\mathbf{X}$.[8] Exogeneity renders (3) a reduced-form relation, or in the terminology of Manski (2013), treatment-response function, which allows for the unbiased estimation of the *total social effect* (Carrell and Hoekstra, 2010; Lavy et al., 2011; Carrell et al., 2018).

However, as Bramoullé et al. (2009) point out, the assumption that all individuals within a group get affected equally by all the members of the group describes a very particular kind of connection structure. We see empirically that, within a group, members often have different interaction intensities; they build social networks that differ 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 $C_r = \{1, \ldots, n_r\}$ that create a graph $(n_r, \mathbf{D}_r)$ through friendships. That is, a network $\mathcal{C}_r$ with vertex set C_r and edge set $D_r = \{(i, j) \in C_r : i \text{ considers } j \text{ to be her friend}\}$.

People usually do not befriend randomly. In fact, socially generated networks share empirical regularities that are incompatible with random formation of links (Jackson and Rogers, 2007).[9] Thus, friendship formation, and in consequence, social networks

[8] See Blume et al. (2011) and Sacerdote (2011) for overviews on the ways to solve identification issues in a reduced-form framework. Examples of more recent reduced-form studies that come to terms with the reflection problem are Lavy et al. (2011) who disentangle endogenous from contextual effects using a classroom survey, Carrell et al. (2013) who create optimally designed peer groups based on causal reduced-form effects in the pre-treatment data and document endogenous reference group formation within the designed groups and Dahl et al. (2014) who use sequential movers as a way to identify endogenous peer peer effects in the take up of parental leave.

[9] First, the average geodesic between pairs of nodes in the network tends to be small. Second, friends

are endogenous (Carrell et al., 2013). Such lack of randomness can occur because of homophily (McPherson et al., 2001; Jackson, 2010; Attanasio et al., 2012; Graham, 2017). 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. Agent i befriends j if there is positive value of doing so: $D_{ij} = \mathbf{1}(Z_{ij}'\beta_Z + U_{ij} > 0)$. Then, homophily implies that dyads in which Z_{ij} is low have higher surplus.[10] We can then write the edge set of network $\mathcal{C}_r$ as $D_r = \{(i,j) \in C_r : Z_{ij}'\beta_Z + U_{ij} > 0\}$. 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.[11] Then, based on Bramoullé et al. (2009) and Goldsmith-Pinkham and Imbens (2013), we extend (2) by incorporating the social interactions collected in network $\mathcal{C}_r$ and operationalized by adjacency matrix $\mathbf{D}$ as mediators of the effects peers have on each agent.

$$\mathbf{Y}_r = \beta_0 \iota_r + \beta \mathbf{X}_r + \beta_{\bar{y}} \mathbf{G}_r \mathbf{Y}_r + \beta_{\bar{x}} \mathbf{G}_r \mathbf{X}_r + \eta_r \tag{4}$$

where $\mathbf{G} = \text{diag}(\mathbf{M})^{-1}\mathbf{D}$ and $\mathbf{M}$ is the vector containing the network's degree sequence (i.e., the total number of friends $M_i = \sum_{j=1}^{n_r} D_{ij}$ each agent has). The reduced-form

tend to cluster into cliques (i.e., linked nodes in the network are tend to have friends in common). Third, compared to networks formed at random, socially formed ones have more nodes with many connections as well as more nodes with fewer connections. Fourth, popular nodes (i.e., those with many connections) are more likely to be linked to other popular nodes, while unpopular nodes are more likely to be linked to other unpopular ones (Jackson and Rogers, 2007).

[10] 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 of 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, 2016; Badev, 2018).

[11] As is convention in the network literature, we assume that one cannot become friends with oneself (Jackson, 2010). That means that $D_{ii} = 0$, $\forall i$.

equation (3) becomes:

$$\mathbf{Y}_r = \frac{\beta_0}{1-\beta_{\bar{y}}}\iota_r + (I_r - \beta_{\bar{y}}\mathbf{G}_r)^{-1}\beta\mathbf{X}_r + (I_r - \beta_{\bar{y}}\mathbf{G}_r)^{-1}\beta_{\bar{x}}\mathbf{G}_r\mathbf{X}_r + (I_r - \beta_{\bar{y}}\mathbf{G}_r)^{-1}\eta_r. \quad (5)$$

Note that unbiased estimation of the total peer-effect $(I_r - \beta_{\bar{y}}\mathbf{G}_r)^{-1}\beta_{\bar{x}}$ requires both $\mathbf{D}$ and $\mathbf{X}$ to be orthogonal to η. However, $\mathbf{D} \not\perp \eta$ because $U_{ij} \not\perp \eta_i$.[12] There are some unobserved components in the link formation process that affect the outcome equation. This confounding problem is the sub-classroom analogoue to the *correlated effects* identification issue (Lin, 2010). We address this issue by modeling the friendship formation process as recent contributions to the literature (Goldsmith-Pinkham and Imbens, 2013; Hsieh and Lee, 2015; Patacchini et al., 2017; König et al., 2018).

Using the surplus structure proposed by Graham (2017), we write $U_{ij} = \theta_{ri} + \theta_{rj} + \xi_{ij}$ and $\eta_{ri} = \theta_{ri} + \nu_{ri}$ as compounded error terms where $Z_{ij} \perp (\theta_i, \theta_j)$ and $\xi_{ij} \perp \nu_i$. Then, if we assume that ξ_{ij} follows a logistic distribution and that they are independently and identically distributed across dyads, we can write the likelihood of observing network $\mathbf{d}$ as

$$\Pr(\mathbf{D} = \mathbf{d}|\mathbf{Z}, \theta) = \prod_{i \neq j} \left[\frac{1}{1+\exp(Z'_{ij}\beta_Z + \theta_i + \theta_j)}\right]^{1-d} \left[\frac{\exp(Z'_{ij}\beta_Z + \theta_i + \theta_j)}{1+\exp(Z'_{ij}\beta_Z + \theta_i + \theta_j)}\right]^{d} \quad (6)$$

Equation (6) implies that conditional on $\mathbf{Z}$ and θ, links form independently. Note that although we allow for $Z_i \not\perp \eta_i$, it is the case that $Z_{ij} \perp \eta_i$. This is the first key identifying assumption of our empirical strategy. It indicates that there are some characteristics that may affect both the link formation and the action choices. However due to homophily, those characteristics will affect the link formation only if matched

[12]Note that $U_{ij} \neq U_{ji}$ because we model directed friendship links

with similar characteristics of another agent at the dyad level. As such, the dyad-level variables are *naturally* excluded through functional form from the outcome equation that models action choices: the difference in dimensions of the dyad-level link formation equation (6) and the individual-level outcome equation (4) forms a natural exclusion restriction. It follows that given the existence of a valid Z_{ij} $\forall i, j \in r$ such that $i \neq j$, we can use (6) not only to instrument $\mathbf{G}$ in the reduced-form (5) and obtain consistent estimates of the total peer-effect, but following Bramoullé et al. (2009) also under certain conditions identify the parameters β_0, β, $\beta_{\bar{y}}$ and $\beta_{\bar{x}}$ of (4).

2.2 Identification through intransitive triads

To outline the identification sources under network endogeneity, let us start by using the series expansion proposed by Bramoullé et al. (2009), where $(I_r - \beta_{\bar{y}}\mathbf{G}_r)^{-1} = \sum_{k=0}^{\infty} \beta_{\bar{y}}^k \mathbf{G}_r^k$, but in our case considering $\mathbf{D}$ to be endogenous, and write (5) as:

$$\mathbf{Y}_r = \frac{\beta_0}{1-\beta_{\bar{y}}}\iota_r + \beta\mathbf{X}_r + (\beta\beta_{\bar{y}} + \beta_{\bar{x}})\sum_{k=0}^{\infty}\beta_{\bar{y}}^k \hat{\mathbf{G}}(\mathbf{Z})_r^{k+1}\mathbf{X}_r + \sum_{k=0}^{\infty}\beta_{\bar{y}}^k \hat{\mathbf{G}}(\mathbf{Z})_r^k \eta_r \tag{7}$$

where $\hat{\mathbf{G}}(\mathbf{Z})_r$ is the notation we use to indicate the matrix resulting from instrumenting the adjacency matrix with (6). Equation (7) implies that, given linear independence of I_r, $\mathbf{G}_r$ and $\mathbf{G}_r^2$ and $\beta\beta_{\bar{y}} + \beta_{\bar{x}} \neq 0$, $(\hat{\mathbf{G}}(\mathbf{Z})_r^2\mathbf{X}_r, \hat{\mathbf{G}}(\mathbf{Z})_r^3\mathbf{X}_r, \dots)$ are the best instruments for $\mathbf{G}_r\mathbf{Y}_r$ in equation (4) and its parameters can be identified using two-stage least squares (2SLS) (Kelejian and Prucha, 1998; Bramoullé et al., 2009). Thus, the presence of intransitive triads (i.e., pairs of i and j that are not connected but share a friend in common) in at least some networks is a sufficient condition to identify the parameters of (4). The strategy thus relies on the exclusion restriction that the characteristics of the agent's friends' friends who are not the agent's direct friends affect her outcomes only through their impact on friends' outcomes (De Giorgi et al., 2010). This is the second key identifying assumption of our empirical strategy. We follow König et al. (2018) and estimate equation (4) by a three-stage least-squares (3SLS) variant of the instrumental variable method of intranstivive triands.[13]

2.3 The importance of distance to a disruptive peer

We can use (7) in a reduced-form strategy to estimate the total peer-effect. However, in the reduced-form application, we use the series expansion as a mechanism to achieve a parsimonious model that provides insights on the mechanisms through which peer-effects take place. We use the fact that $\beta_{\bar{y}}^{k}\mathbf{G}_r^{k+1}$ is simply discounting by a parameter $\beta_{\bar{y}}^{k}$ the effect that neighbors k-steps apart can have on the agent. In consequence, $(\beta_x\beta_{\bar{y}}+\beta_{\bar{x}})\sum_{k=0}^{\infty}\beta_{\bar{y}}^{k}\mathbf{G}_r^{k+1}\mathbf{X}_r \approx \mathbf{\Gamma}(\|\mathbf{r}\|)\mathbf{X}_r$, where $\mathbf{\Gamma}(\cdot)$ is a function whose argument $\|\mathbf{r}\|$ denotes a matrix with the geodesic between the nodes in network r. The equation to estimate becomes

$$\mathbf{Y}_r = \phi_0\iota_r + \phi_1\mathbf{X}_r + \Gamma(\|\mathbf{r}\|)\mathbf{X}_r + \varepsilon_r \tag{8}$$

where we instrument $\|\mathbf{r}\|$ with the corresponding propensity of friendship based on homophily given by (6). We take advantage of the fact that (6) provides several instruments for every element in $\|\mathbf{r}\|$, say the geodesic between i and j, because one can not only include the $\Pr(D_{ij}=1|Z_{ij},\theta_i,\theta_j)$ but also $\Pr(D_{il}=1|D_{ij}=0,D_{jl}=1,Z_{il},\theta_i,\theta_l)$, $\Pr(D_{hj}=1|D_{ij}=0,D_{ih}=1,Z_{hj},\theta_h,\theta_j)$ and so on. The estimation of (8) is informative

[13] An alternative approach to the instrumental variables approach adopted here is a selection model similar to the control function approach of Heckman (1979) and Heckman and Robb (1986). This control function approach was proposed by Blume et al. (2015) and has been adopted in the linear-in-means model context by Goldsmith-Pinkham and Imbens (2013) and Hsieh and Lee (2015) who estimate the model with fully parametric specifications of both the outcome and dyadic link formation equations using Bayesian methods. Two recent studies propose frequentist solutions to estimating the linear-in-means model using a control function approach (Johnsson and Moon, 2017; Auerbach, 2019). Both studies model dyadic link formation nonparametrically.

because $\nabla \mathbf{y}_{ir} / \nabla \mathbf{X}_r = \phi_1 \mathbf{e} + \Gamma(\|\mathbf{r}\|)$, 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 potentially depend on the path length between the two agents, i.e., how socially close they are.

We impose no restriction on $\mathbf{\Gamma}(\cdot)$ except for the fact that it should fulfill the rank condition with respect to the number of instruments available from (6). We favor parsimonious choices of $\mathbf{\Gamma}(\cdot)$, but it could even be non parametric. We recognize that parsimonious choices of $\mathbf{\Gamma}(\cdot)$ may render the approximation to $(\beta_x \beta_{\bar{y}} + \beta_{\bar{x}}) \sum_{k=0}^{\infty} \beta_{\bar{y}}^k \mathbf{G}_r^{k+1} \mathbf{X}_r$ less precise especially for large k. However, literature has established that the average geodesic between pairs of agents and the diameter (the maximum distance between any pair of nodes) of a social network are usually small (Jackson and Rogers, 2007). Hence, in our reduced-form empirical strategy we implement $\mathbf{\Gamma}(\|\mathbf{r}\|) = \gamma_1 \iota_r \|\mathbf{r}_{k \leq \kappa}\| + \gamma_2 \iota_r \mathbf{1}[\|\mathbf{r}_{k > \kappa}\|]$, where κ is a given distance after which we assume that further away neighbors will contribute in the same amount to the effect. See Section 4 for further details.

3 Data: The Stockholm Birth Cohort

The Stockholm Birth Cohort Study (SBC) follows all the individuals that were born in 1953, were living in the Stockholm metropolitan area in 1963 and were alive and living in Stockholm in 1980 and/or 1990. The original sample of 1953 identified in 1963 comprised 15,117 people, of which 14,950 survived until 1980 and were eligible for matching to contemporary Swedish register data. In total, 14,294 (96 percent) within 18 metropolitan municipalities were identified and linked to registries from 1980 onwards.[14]

The SBC combines a number of registries and surveys covering the cohort members' early life and adolescence. It contains extensive information on individual characteristics including prenatal and perinatal information, cognitive test scores, school performance, and educational attainment. In addition, the SBC collected information on social circumstances and family background such as household composition, parents' education and occupation, the neighborhood block they were living in at the time when the child was born, and parental quality. Table 1 present summary statistics of the variables used in this paper.

We drop classrooms with less than 7 students as they presumably hold very different social networks to the one observed in regular sized classrooms. We also drop 56 schools that have only one classroom per grade, and therefore are not subject to the within-school variation we exploit in our empirical strategy, and 28 schools with special education classes. We end up with an analytic sample of 7,826 students from 387 classes belonging to 118 schools.[15]

3.1 Measures

The SBC contains the following information necessary for our analysis.

Disruptiveness. We proxy disruptive peers with children whose parents were considered abusive and neglectful by the CWC at some point between ages 0 to 12 years.

[14]See Stenberg and Vagerö (2006) for a cohort profile and www.stockholmbirthcohort.se for codebooks for the included registries.

[15]We use the direct individual-level survey responses on attendance of a special education class. In total 273 students attended special education classes. We further validate that these special education classes encompass all deviant classrooms in the data by conducting a cluster analysis based on class averages of student's intelligence tests in sixth grade.

Table 1: Summary Statistics

Variable	Mean	Std. Dev.	N
Disruptive peers (DP)			
Students per Classroom	26.4	3.6	387
Share of Classrooms with a DP	0.618	0.487	387
No of DP in the Classroom	1.124	1.247	387
No of DP $\vert$(DP in Classroom$\geq$1)	1.820	1.118	239
Share of DP in the Classroom	0.054	0.058	7,826
Leave-out Mean Ratio of DP	0.054	0.059	7,826
Share of DP $\vert$ (DP in Classroom$\geq$1)	0.087	0.052	4,864
Controls			
Female	0.504	0.5	12,047
Parents Social Aid Recipients†	0.133	0.34	12,047
Birthweight (kg)	3.5196	0.488	12,047
Mother's Age	28.699	5.757	12,047
Own Dwelling	0.183	0.387	11,507
Dwelling Size	0.974	0.16	11,499
Older Siblings	0.884	1.021	12,039
Outcomes			
Verbal	25.185	6.358	11,434
Numeric	21.247	7.807	11,444
Spatial	23.083	6.928	11,431
Marks in Grade 9 *100	318.839	76.658	11,413
Verbal (males age 19)	0	1.299	5,402
Spatial (males age 19)	6.111	1.711	5,416

Note: The summary statistics for disruptive peers within classrooms refer to the analytic sample (see Section 3 for sample restrictions). The other descriptive statistics (outcomes and controls) are reported for the full sample. Data from Stockholm Birth Cohort.† indicates a binary variable.

The CWC investigated these parents on the suspicion of abuse and neglect and decided to take measures to protect the child. We only consider the investigations that resulted in measures (some 10 percent of the investigations were dismissed). For all students in the SBC, for which we observe classroom data (n=12,047), we observed 831 CWC investigations that match our definition, of which 655 (78.8 percent) were carried out before the child started school. In our analytic sample, we observe 435 children whose parents were at least once investigated for abuse and neglect. Of these, 379 children's parents were investigated before the child started school. Of all the 435 investigations in our analytic sample 381 cases led to child removal and placement in foster care or institution. According to our data, the main reasons for initiating an investigation for abuse and neglect were parents' alcohol abuse, parents' psychiatric disorders and parent's death.[16]

Social interactions. Friendship nominations were measured in a classroom survey conducted for sixth graders (age 13). Each student was asked to nominate her three best friends in the classroom. Of all students in our data who participated in the school survey (n=11,444) , 6,785 nominated three friends (59.3 percent), 3,191 nominated two friends (27.9 percent), 906 nominated only one friend (7.9 percent) and 557 did not nominate any friends (4.9 percent). In our analytic sample (n=7,045), 4,227 students nominated three friends (60.0 percent), 1,940 students nominated two friends (27.5 percent), 524 students nominated only 1 friend (7.5 percent) and 354 students did not nominate any friends (5.0 percent).

[16] Presumably, another reason for initiating an investigation could have been parents' involvement in criminal activities, however the CWC files do not record crime records of parents.

Prenatal and perinatal health. The SBC contains the delivery records of the cohort members. They are administrative birth registries that collect health information about the mother, the newborn child, the pregnancy and the delivery, as well as the health of past pregnancies. In particular, regarding the mother's health condition and practices, we use the reports of illness during pregnancy (e.g., toxemia, fever, anemia), if the mother had prenatal visits with a physician, the mother's age at the time of the child's birth, the length of stay in the hospital after giving birth, whether this was her first pregnancy and whether she had previous children who died. Regarding the delivery, we use information on whether it was a vaginal delivery, it required the use of forceps, or the use of a c-section. Finally, regarding the newborn's health, we use birthweight and the height of the baby at birth. In some specifications, we summarize the information of all these health indicators with the use of the first two principal components.

Cognitive tests. We use the three components of ability collected by the school survey in 1966 (at age 13): numeric, verbal and spatial ability. Numeric ability was tested by presenting 40 numerical sequences of six numbers each of which follow a logical pattern based on elementary arithmetic concepts. The students were asked to predict the next two numbers in the sequence following the same pattern. Verbal ability was tested by presenting the respondent with 40 words for which the student had to find antonyms among four options per each given word. The spatial ability test consisted of 40 figures which are unfolded and need to be folded mentally. The verbal and numeric tests are weighted more toward crystallized intelligence while the spatial cognitive ability test is weighted more towards fluid intelligence which is often considered to be the purer of the two measures of 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). In fact, the numeric and verbal tests might even more appropriately be called achievement tests than intelligence test (Almlund et al., 2011). Tests of fluid intelligence, such as spatial ability tests, are conventionally designed to capture innate ability clean from acquired knowledge.

Marks in ninth grade (Marks 9) refers to the Spring term grade point average on a grade scale from 1 to 5 (the variable is rescaled by multiplying by 100). Verbal and spatial ability at age 19 come from a mental ability test collected for males in enlistment records filed by the National Conscript Board. Verbal ability in this context measures ability to grasp quickly, inductive ability and verbal comprehension. Spatial ability was measured by an assembling test including 25 items. Both tests were measured using the stanine score scale (1-9).

4 Empirical strategy

Section 2 derives three important models for us to test empirically with our data. The next few subsections outline the empirical strategies we use to estimate the reduced-form total peer-effect of having disruptive classmates, the importance of social distance to the disruptive classmate, and the instrumental variable model identifying the structural parameters of the linear-in-means model through intransitive triads.[17]

[17] See Blume et al. (2011) for an extensive overview of the various approaches to identify the linear-in-means model, both in its reduced and structural forms.

4.1 Negative externalities: The reduced-form of the linear-in-means model

In this subsection we address to what extent having disruptive peers affect the student's cognitive outcomes by estimating the reduced-form model of equation (3). To identify the total peer-effect parameter, we follow the standard protocol of a reduced-form specification of a Manski type linear-in-means model. Our identification strategy sorts out a total (be it endogenous or contextual) negative peer effect from correlated effects (independent of any social interaction) through proxying troubled peers by parents who were investigated for abuse and neglect and exploiting idiosyncratic differences in the composition of individual classes within a school (Carrell and Hoekstra, 2010). We argue that this is an exogenous measure 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 feeback 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). Table 2 shows that we find no correlation between own exposure to abuse and neglect and share of peers who have experienced abuse and neglect.[18]

We observe all classrooms in sixth grade of upper primary school within every school. Since there was no tracking of students to upper primary school, students were assigned to classes within schools as good as randomly with respect to family background. 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

[18] To address the mechanical negative bias inherent to this test originally raised by (Guryan et al., 2009), we control for the complete set of peers exposed to abuse and neglect within ones school that could have ended up in one's own classroom.

Table 2: Exposure to Disruptive Peers and Own CWC Condition

	Own CWC Decision	
	Coef.	*Std.Err.*
% Male Peers with CWC Decision	0.065	(0.052)
% Female Peers with CWC Decision	0.033	(0.037)
Observations	10,213	

Note: Data from Stockholm Birth Cohort. The regression includes school fixed effects, exclude classes with less than 7 students and schools with special education classrooms. Robust standard errors in parentheses clustered at the school level. Following Guryan et al. (2009). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

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 be exposed to 6 percent of students with abusive parents. In Table 3, we document that the share of troubled peers in the classroom is uncorrelated with a number of important background characteristics while controlling for school fixed effects.[19] This evidence substantially reduces the concern that correlation in unobserved background characteristics affecting the decision to join the reference group would confound the causal effect of the reference group mean outcome on the individual outcome.

Our strategy is facilitated by the institutional framework that governed assignment to Swedish schools in the 1960's. Back then, students attended a single-tracked compulsory school with six years of primary school and three years of upper secondary school. Primary school was divided into two stages: grades 1 to 3 of lower primary school

[19]Table B.1 of the Appendix reports the results of a similar balancing test that a binary variable of whether the classroom had at least one student whose parents were investigated for abuse and neglect.

Table 3: Exposure to Disruptive Peers and Own Characteristics

	Gender		Social Aid Recipt		Birth Weight		Mother Age	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
% of Peers with CWC evaluation								
All	0.000	-0.001	0.028***	0.000	-0.032**	-0.005	0.004**	0.001
	(0.002)	(0.001)	(0.007)	(0.002)	(0.018)	(0.013)	(0.003)	(0.001)
Male	0.001	0.000	0.012***	0.000	-0.005	0.002	0.002	0.000
	(0.001)	(0.001)	(0.003)	(0.001)	(0.012)	(0.010)	(0.001)	(0.001)
Female	-0.001	-0.001	0.015***	0.000	-0.027***	-0.007	0.002*	0.000
	(0.001)	(0.001)	(0.005)	(0.002)	(0.011)	(0.009)	(0.002)	(0.001)
Schoool FE	N	Y	N	Y	N	Y	N	Y
Observations	7,459	7,459	7,459	7,459	6,322	6,322	6,486	6,486

	Owner of Dwelling		Dwelling Size		Older Siblings	
	(1)	(2)	(1)	(2)	(1)	(2)
% of Peers with CWC evaluation						
All	-0.023***	0.001	-0.016***	-0.005	0.000	-0.001
	(0.005)	(0.002)	(0.007)	(0.004)	(0.001)	(0.001)
Male	-0.009***	0.001	-0.006*	-0.002	-0.001	-0.001
	(0.003)	(0.002)	(0.004)	(0.003)	(0.001)	(0.000)
Female	-0.015***	0.000	-0.011**	-0.003	0.000	0.000
	(0.003)	(0.001)	(0.005)	(0.003)	(0.001)	(0.000)
Schoool FE	N	Y	N	Y	N	Y
Observations	7,459	7,459	7,459	7,459	7,459	7,459

Note: Data from Stockholm Birth Cohort. Each cell reports results from a separate regression. All regressions exclude classes with less than 7 students and schools with special education classrooms. Sample restricted to individuals who were not intervened by CWC. Robust standard errors in parentheses clustered at the school level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

(from age 7 to 10) and upper primary school up until completion of sixth grade (at age 13). No tracking occurred at the transition from lower primary school to upper primary school.[20] Compulsory school funding was centralized by the government, and school assignment was determined by place of residence.[21]

We reparametrize (3) as

$$\mathbf{y} = \lambda_0 + \lambda_1 \mathbf{X} + \lambda_2 \mathbf{J}_r \mathbf{X} + \lambda_3 \mathbf{C} + e. \tag{9}$$

where $\mathbf{X}$ is a dummy variable indicating whether the student's parents at some point during her childhood were investigated by CWC for abuse and neglect. As in (3), projection matrix $\mathbf{J}_r = (n_r - 1)^{-1}(\iota_{n_r}\iota'_{n_r} - I_{n_r})$ represents the leave-out mean operator. Estimates relating individual outcomes to leave-out means of peer characteristics are prone to bias due to a *mechanical link* between own characteristics and peer characteristics (Angrist, 2014). To ameliorate this issue, we follow Angrist's recommendation and separate between those who are subject to the peer effects and the peers who are hypothesized to provide the mechanism of causal effects on these subjects and drop the abused and neglected students for whom $\mathbf{X}$ takes on value 1. This renders the following estimating equation:

$$\mathbf{H}_s \mathbf{y} = \omega_1 \mathbf{H}_s \mathbf{J}_r \mathbf{X} + \mathbf{H}_s \mathbf{C} \phi + \mathbf{H}_s \upsilon \tag{10}$$

[20] Formal grading of students only began in seventh grade which substantially mitigates concerns of any informal tracking of students based on ability.

[21] The compulsory school system had several organizational layers. The school formed the principal unit that was aggregated up to school districts that normally consisted of one lower secondary school and one or many primary schools. Here we follow the terminology used by Fredriksson et al. (2012). In fact, 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 was 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 school in one district.

where the projection matrix $\mathbf{H}_s = (I_{n_s} - \frac{1}{n_s}\iota_{n_s}\iota'_{n_s})$ transforms the variables into demeaned form in order to eliminate school fixed effects and $\mathbf{C}$ is a vector of additional predetermined controls. The reduced-form parameter of interest ω_1 measures the existence of a social effect of troubled peers on student i's outcome Y.

4.2 Endogenizing friendship formation

The next step of our empirical approach is to incorporate the intensity of social interactions as mediator of the peer effects. To do so, we acknowledge that social link formation is endogenous. Friendship is an active choice and, in consequence, the shape of a given network is going to be the result of those choices. We address this issue relying on homophily. As indicated in Section 2, under homophily, the probability of a friendship between i and j decreases with the distance in the covariate space $Z_{ij} = |Z_i - Z_j|$. Thus, we estimate equation (6), the probability of that i considers j to be her friend , with the following logistic regression function allowing for degree heterogeneity:

$$\Pr(D_{ij} = 1|\mathbf{Z}, \theta) = \frac{\exp(\alpha_0 + \alpha_z Z_{ij} + \theta_i + \theta_j)}{1 + \exp(\alpha_0 + \alpha_z Z_{ij} + \theta_i + \theta_j)} \tag{11}$$

where the 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,[22] 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

[22] We use the finest level of neighborhood recorded in Stockholm Birth Cohorts which is defined as district of residence in the codebook (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 473 neighborhoods.

indices that summarize their prenatal and perinatal health as well as the health of their mothers during their pregnancies as evidenced by the birth records in 1953 (i.e., Neonatal health 1 and Neonatal health 2). The indices are the first two principal components of the system of measures summarized under the title *Prenatal and perinatal health* in Section 3.[23]

The choice of the dyadic variables are not based on economic theory. Our goal is to achieve the best possible fit of the dyadic link formation model based on predetermined characteristics rather than to explore the causal channels through which observed homophily materializes. May it suffice to say that previous evidence shows homophily to be a widely observed feature of social networks (McPherson et al., 2001; Jackson, 2010), and that previous studies have modeled link formation between individuals i and j based on the utility derived from similarity in observed and unobserved characteristics (Goldsmith-Pinkham and Imbens, 2013; Hsieh and Lee, 2015; Graham, 2017; Patacchini et al., 2017). Our proposed predetermined variables are proxying for similarity in socioeconomic background, health endowments and spatial proximity, although importantly, in characteristics that are arguably excluded at the dyad-level from the outcome equation. Hence, the identification comes from exclusion restrictions through functional form (Hsieh and Lee, 2015).

While modeling link formation as in equation (11) we abstract from interdependencies in agents' decision to form a link. The implicit assumption underlying our model is that agents i and j may form a link regardless of whether i and k, where $k \neq j$, have formed a link or not and as such the model has a unique equilibrium (Sheng, 2016; Badev, 2018).[24] Graham (2015) notes that a fixed effects model of degree heterogeneity

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

[24] Graham (2017) discusses the plausibility of this assumption in different contexts. He proposes a

is likely to predict the link probability well even if the true link formation process happened to include strategic aspects.

We estimate equation (11) using a Logit model with sender (i.e., individual i) fixed-effects, while controlling for observed variation in the following characteristics of the receiver of the friendship nomination (i.e., friend j): gender, social aid and the indices Neonatal health 1 and Neonatal health 2. A joint fixed effects error term structure $\theta_j + \theta_j$ as proposed by Graham (2017) would require undirected links. Given that students nominated their friends through individual surveys and that we know exactly who nominated whom in reference group r, we consider friendships to be asymmetric affairs. Hence we model directed links (Lin, 2010; Liu et al., 2014; Hsieh and Lee, 2015; Charbonneau, 2017); if student i nominated student j but not vice versa, then entry g_{ij} of the weighting matrix will have a positive value and entry g_{ji} will be zero. See the regression results of estimating equation 11 in Table 5. We use these estimates to predict friendship links.

4.3 All peers don't matter equally

In order to further understand how social connections moderate the negative externality of disruptive peers, we exploit information on social interactions and relax the assumption made that peers affect everyone in the classroom equally. As indicated in Section 2.3, and in paritcular in equation (8), we model the effect of social closeness to the disruptive peer on cognitive outcomes based on the geodesic (number of edges

way to incorporate interdependent preferences in the link formation model with agent-level heterogeneity using panel data on the networks. However, we only observe the network once. As mentioned before, an additional wrinkle in the presence of strategic interdependencies would be that the researcher is confronted with the possibility of multiple equilibria of the network formation model (Sheng, 2016; Badev, 2018). For further discussion and references on modeling strategic interdependencies, see (Graham, 2015; de Paula, 2017).

separating the two individuals within the social network) of each student in the classroom to the troubled peer, $\|\mathbf{r}_{k\leq\kappa}\|$, where κ is a given number of edges after which we assume that more distant peers will have a constant effect on the student's outcomes, modeled by an indicator function $\mathbf{1}[\|\mathbf{r}_{k>\kappa}\|]$. In practice, κ is defined by the social network formed by the students' active choices of peers. This link formation process may leave several students within the classroom unconnected with each other. In our data, we observe several classes comprising two or more disjoint social networks resulting in students who, although being classmates, are not connected socially. This allows us to inquire about the mechanism through which the effect of having abused and neglected peers materializes. If the circumstance of having a troubled peer in the classroom is enough to disrupt learning, we expect to find an effect even among those who are not socially connected to her. If on the contrary, behavior and its direct impact through social connections is the main channel of peer effects, we should not observe much of an effect among those who share a classroom with her, but are socially far or detached from her.

Instead of considering the leave-out fraction of troubled peers, $\mathbf{X}_{rs}$, we consider an indicator that takes on value one if there is at least one troubled peer in the classroom, $\mathbf{1}[\iota_r\mathbf{X}_{rs} > 0]$.[25] The endogeneity of friendship formation renders both $\mathbf{1}[\|\mathbf{r}_{k>\kappa}\|]$ and $\|\mathbf{r}_{k\leq\kappa}\|$ endogenous in our reduced-form model (8). In the spirit of König et al. (2018), we propose a solution by instrumenting $\|\mathbf{r}\|$ with a set of propensities of friendship estimated in (11).

[25] We show in Appendix B that the balancing test also goes through when exposure to disruptive peers in the classroom is defined as a binary variable.

Our peer-effect model then becomes:

$$\mathbf{y}_{rs} = \beta_0 \iota_{rs} + \beta_{\mathbf{1}[.]} \mathbf{1}[\iota_{rs} \mathbf{X}_{rs} \geq 0] + \gamma_1 \iota_{rs} \|\mathbf{r}_{k \leq \kappa}\| + \gamma_2 \iota_{rs} \mathbf{1}[\|\mathbf{r}_{k > \kappa}\|] + \mu_s + \varepsilon_{rs} \tag{12}$$

using the probability that student i forms a link with the disruptive peer q, $\hat{\Pr}(D_{iq} = 1|Z_{iq}, \theta_i, \theta_q)$, as an instrument for the geodesic distance between i and q, $\|\mathbf{r}_{k \leq \kappa}\|$, and both $\hat{\Pr}(D_{iq} = 1|Z_{iq}, \theta_i, \theta_q)$ and $\hat{\Pr}(D_{lq} = 1|D_{iq} = 0, Z_{lq}, \theta_l, \theta_q)$ as instruments for the indicator function, $\mathbf{1}[\|\mathbf{r}_{k > \kappa}\|]$, attaining value 1 if individual i is not connected to q. This IV solution forms a system of three equations that we estimate using Limited Information Maximum Likelihood estimation (LIML).

4.4 The structural model

We estimate the structural model (4) using a 3SLS model similar to König et al. (2018) in which the first step deals with the endogenity concern of *correlated effects* resulting from endogenous network formation by estimating our link formation model of equation (6) operationalized in (11). We obtain *all* the predicted links $\hat{D}(Z)_{ij}$, in contrast to the estimation strategy in Section 4.3 where only the distances to the troubled peer were collected. Then after replacing the endogenous row-normalized adjacency matrix $\mathbf{G}$ by the *predicted* row-normalized adjacency matrix $\hat{\mathbf{G}}(\mathbf{Z})$, we use a 2SLS, using the identifying power of intransitive triads (i.e., matrices $\mathbf{I}$, $\mathbf{G}$ and $\mathbf{G}^2$ are linearly independent) in order to deal with the endogeneity concern of the reflection problem (Bramoullé et al., 2009). We incorporate school fixed effects, in order to leverage on the random sorting of troubled peers across classrooms:

$$\mathbf{H}_s \mathbf{y}_{rs} = \beta \mathbf{H}_s \mathbf{X}_{rs} + \beta_{\bar{y}} \mathbf{H}_s \mathbf{G}_{rs} \mathbf{Y}_{rs} + \beta_{\bar{x}} \mathbf{H}_s \mathbf{G}_{rs} \mathbf{X}_{rs} + \mathbf{H}_s \eta_{rs} \tag{13}$$

In our case, Bramoullé et al. (2009) 2SLS simplifies to a standard 2SLS as we account for all the spacial correlation explicitly through $\bar{Y}$ and $\bar{X}$.[26] The matrix of explanatory variables $\widetilde{\mathbf{X}} = [\mathbf{Gy} \quad \mathbf{X} \quad \mathbf{GX}]$ is instrumented in the second stage by $\mathbf{S} = [\mathbf{X} \quad \hat{\mathbf{G}}(\mathbf{Z})\mathbf{X} \quad \hat{\mathbf{G}}(\mathbf{Z})^2\mathbf{X} \quad \hat{\mathbf{G}}(\mathbf{Z})^3\mathbf{X}]$ yielding in the third, and final, stage the estimator $\hat{\theta}^{2SLS} = (\widetilde{\mathbf{X}}'\mathbf{P}\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{X}}'\mathbf{Py}$.[27] Here, $\mathbf{P}$ is the projection matrix $\mathbf{S}(\mathbf{S}'\mathbf{S})^{-1}\mathbf{S}$.

5 Results

5.1 Externalities of abused and neglected peers on cognitive scores and educational attainment

Table 4 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 marks in ninth grade. In general, the inclusion of one additional disruptive classmate to a classroom of 20 students decreases verbal and numeric ability in sixth grades by 0.20 (=0.05*4.033) and 0.24 (=0.05*4.747) points respectively. The estimates amount to negative changes of 3.2 percent and 3.0 percent of a standard deviation in their respective scales. The average detrimental effect on cognitive achievement is larger among girls than among boys. Our estimates imply that adding one more disruptive peer to a classroom of 20 students decreases girls' verbal and numeric ability scores by 4.2 percent and 4.3 percent of a standard deviation respectively. The fourth column in Table 4 indicates that these effects on

[26]The Generalized 2SLS strategy first proposed by Kelejian and Prucha (1997, 1998) and adopted by Bramoullé et al. (2009) reduces to standard 2SLS whenever disturbances are not spatially correlated.

[27]The within-transformation matrix $\mathbf{H}_s$ is retained throughout the analysis.

girls' cognitive ability linger enough to affect their school grades three years later by 4.2 percent of a standard deviation.

When examining the heterogeneity of these results by gender of the peer exposed to abuse and neglect, we find that while disruptive boys affect their male and female classmates, disruptive girls have negative externalities on other girls but not on boys. For instance, the inclusion of an additional disruptive *boy* to a classroom of 20 students decreases other boys' numeric ability by 6.7 percent of a standard deviation and girls' numeric ability by 3.9 percent of a standard deviation. This contrasts with the effect of the inclusion of an additional disruptive *girl* to a similar classroom, which will have no significant effect on boys' numeric ability, but would have a negative effect of 4.8 percent of a standard deviation on other girls' numeric ability.

The effect disruptive boys have on other boys' cognitive ability extends into adulthood. Our estimates indicate that having had one additional disruptive boy in a classroom of 20 students in elementary school decreases verbal ability as measured at age 19 in the enlistment records (for men only) by 5.9 percent of a standard deviation.

We report throughout the analysis results for fluid intelligence as measured by the spatial ability component of the intelligence test. Given that fluid intelligence is considered to be innate and much less responsive to external inputs, we find it reassuring that we estimate no significant peer effects on spatial ability test scores. Combined with the significant results on test scores capturing crystallized ability, this null effect could potentially be interpreted as evidence supporting our identification strategy in the spirit of a placebo test.

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 Florida. Our estimates, ranging between 3 percent and 4 percent of a standard deviation, are in the same order of magnitude as what they document (i.e., 2.5 percent of a standard deviation).

5.2 Homophily

As mentioned in Section 4.3, one of our contributions is that we do not take adjacency matrix $\mathbf{D}$ as exogenous. We endogenize social connections modeling them in (11) as products of homophily and degree heterogeneity. Table 5 shows strong evidence in favor of the existence of homophily. Children are more likely to befriend classmates from the same gender and the same socioeconomic condition as measured by being or not a welfare recipient, home owners or having father with the same level of occupation at the time of birth. The probability of friendship also increases with geographical proximity. Children are more likely to be friends of classmates whose families lived in the same neighborhood block at the time of their birth. This may not only be proxying for more detailed dimensions of the socioeconomic condition, but also for parental social relations and even for the fact that they may go (come) to (from) school together. Table 5 also shows that children are more likely to befriend classmates who had similar prenatal and perinatal health. Namely, the probability of friendship increases if both children share their birth order, or if the difference between their prenatal and perinatal health scores is small.

Table 5 shows four different models. They differ in the type of geographic detail we use (for Address53B we aggregated some scarcely populated suburban residential blocks) and whether or not we control for some individual characteristics of the friendship

Table 4: Effects of Disruptive Peers on Cognitive Scores at Ages 13, 16 and 19

	IQ Components at 13			*Grades at 16*	*Tests at 19*	
	Crystallized Int.		Fluid Int.			
	Verbal	Numeric	Spatial	Marks 9	Verbal	Spatial
% CWC Peers	-4.033*	-4.747*	-2.034	-37.746*		
	(2.169)	(2.635)	(2.653)	(21.694)		
Observations	7,032	7,032	7,032	6,751		
% CWC Peers						
Men	-2.671	-3.224	-0.860	-9.701	-0.690	0.448
	(2.406)	(3.039)	(3.317)	(27.844)	(0.509)	(0.591)
Women	-5.385**	-6.644**	-3.632	-63.777***		
	(2.539)	(3.088)	(2.547)	(24.333)		
Observations	7,032	7,032	7,032	6,751	3,156	3,164
% Male CWC Peers						
Men	-6.540*	-10.508**	-1.685	-31.670	-1.538**	-0.207
	(3.787)	(4.842)	(5.124)	(44.499)	(0.725)	(0.779)
Women	-6.104*	-6.149	-5.101	-76.938**		
	(3.302)	(4.677)	(3.841)	(36.838)		
% Female CWC Peers						
Men	0.714	2.943	0.130	13.486	-0.014	1.009
	(3.059)	(4.252)	(4.148)	(38.906)	(0.687)	(0.775)
Women	-5.125	-7.520*	-2.712	-59.199*		
	(3.822)	(4.037)	(3.579)	(31.942)		
Observations	7,032	7,032	7,032	6,751	3,156	3,164

Note: Data from Stockholm Birth Cohort. Independent variables are the percentage of peers with CWC decision and its interaction with female. Sample restricted to individuals who were not intervened by CWC. Sample also excludes classes with less than 7 students and schools with special education classrooms. All regressions include school fixed effects and control for gender, 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. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

nomination's potential receiver. Results are remarkably stable across specifications. Given that we want the best prediction possible, we will construct the predicted links $\hat{d}_{ij}$ from the estimates in column (4), which contain slightly more information.

Table 5: Probability of Forming a Friendship Link Between Two Classmates

	(1) $d_{ij}=1$	(2) $d_{ij}=1$	(3) $d_{ij}=1$	(4) $d_{ij}=1$
$\mathbf{1}[x_i = x_j]$				
DeadSiblings	-0.014	-0.014	-0.016	-0.016
	(0.027)	(0.027)	(0.027)	(0.027)
First Child	0.056***	0.056***	0.059***	0.059***
	(0.021)	(0.021)	(0.021)	(0.021)
PrenatCareDr	0.029	0.029	0.040*	0.039*
	(0.022)	(0.022)	(0.022)	(0.022)
Gender	3.739***	3.739***	3.740***	3.740***
	(0.044)	(0.044)	(0.044)	(0.044)
Social Aid	0.256***	0.257***	0.211***	0.211***
	(0.028)	(0.028)	(0.036)	(0.036)
Own	0.135***	0.133***	0.136***	0.134***
	(0.028)	(0.028)	(0.028)	(0.028)
FatherClass53	0.053***	0.053***	0.052***	0.053***
	(0.020)	(0.020)	(0.020)	(0.020)
Address53	0.172***		0.174***	
	(0.033)		(0.033)	
Address53B		0.173***		0.175***
		(0.035)		(0.035)
$\lvert x_i - x_j \rvert$				
Natal Health 1	-0.000	-0.000	-0.000	-0.000
	(0.002)	(0.002)	(0.002)	(0.002)
Natal Health 2	-0.007***	-0.007***	-0.007***	-0.007***
	(0.002)	(0.002)	(0.002)	(0.002)
x_j				
Female			-0.037	-0.037
			(0.044)	(0.044)
Natal Health 1			-0.004	-0.004
			(0.007)	(0.007)
Natal Health 2			-0.025***	-0.025***
			(0.008)	(0.008)
Social Aid			-0.066*	-0.067*
			(0.036)	(0.036)

Note: Data from Stockholm Birth Cohort. Sample excludes classes with less than 7 students and schools with special education classrooms. We model directed links. The dependent variable d_{ij} takes on value 1 if i nominated j (and zero otherwise, even in the event that j nominated i, i.e., $d_{ji} = 1$). The number of observations (i.e., potential links) in all regression is 157,298 that stem from a sample of 7,506 unique students. We estimate fixed-effect logit regressions. We include agent i (i.e., sender) fixed-effects. *Address53* is the individuals neighborhood of residence at birth at the precision of block in many cases. The students in the analytic sample lived in 1953 in 862 different neighborhoods. *Address53B* replaces *Address53* with neighborhood or parish in 1963 for those who were born in a neighborhood located outside the Stockholm metropolitan area (and moved in before 1963 so as to be included in the sampling frame). The students in the analytic sample lived in 473 such neighborhoods located within the Stockholm metropolitan area. We include dummies to control for missing values at the students level. Their coefficients are provided upon request. Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

5.3 Distance

In this subsection, we present the estimates of the model presented in equation (12). Its operationalization includes a dummy of having a disruptive peer in classroom; a dummy that takes the value of one if, in the presence of a disruptive peer in the classroom, the student is not socially connected to her; and the geodesic to the closest disruptive peer to which a student is socially connected. The social connectedness variable and the geodesic to the closest disruptive peer are endogenous and thus we model them jointly with the outcome equation based on the homophily results shown above.

Results in Table 6 indicate that social networks are important in determining the size of the negative effect of having a disruptive peer in the classroom. The direct friends of those peers (i.e., those with distance equal to one) suffer the greatest negative effects. Namely, the verbal ability and numeric ability test scores decrease by -0.795 and -0.995 points respectively. This implies a reduction in verbal test scores by 12.5 percent of a standard deviation and a reduction in numeric test scores by 12.7 percent of a standard deviation. These effects of having a direct friends who is disruptive are substantially larger than the estimated total peer-effect of an additional disruptive peer into a classroom of 20 students of roughly 3 percent of a standard deviation in test scores reported in subsection 5.1. As the geodesic to the disruptive peer increases, the negative effect dissipates following an inverse distance rule.[28] For instance, friends of friends of the disruptive peer suffer a decrease of -0.394 points in the verbal score and -0.360 points in the numeric score. Our results show that those who are three edges away from the neglected peer incur almost no harm of being associated with her socially.

[28]The inverse distance rule has previously been found in students' generosity with their peers. Experimental evidence shows students choose to share a larger proportion of the payoff with those socially closer to them (Leider et al., 2009; Goeree et al., 2010)

The importance of social interactions in communicating peer effects are also evident among those not connected. Our results show that not being socially connected to the disruptive peer, as measured by directed links, offsets almost entirely the negative effect of having her in the classroom. Table 6 also shows effect heterogeneity of social interactions across the gender subgroups. While direct boy peers of the disruptive peers suffer a -0.924 point decrease in the verbal score the effect is substantially much smaller for girls (-0.556) and insignificant. For girls overall, the results are less precise than for boys.

As discussed in Subsection 5.1, spatial ability test scores could potentially be thought of as a placebo outcome in our analyses. The estimates in Table 6 repeat the same pattern as the reduced-form estimates of total peer-effects reported in Table 4; compared to the effect of distance to the disruptive peer on numeric and verbal ability test scores, results on spatial ability test scores show a rather different pattern and the coefficients are insignificant in almost all models.

5.4 Results of the structural model

In Table 7, we present the estimation results of equation (13). We present for each outcome two different models. In the first column for each outcome, we use as excluded instruments for $\mathbf{Gy}$ both $\hat{\mathbf{G}}(\mathbf{Z})^2\mathbf{X}$ and $\hat{\mathbf{G}}(\mathbf{Z})^3\mathbf{X}$. In the second column, we use only $\hat{\mathbf{G}}(\mathbf{Z})^3\mathbf{X}$ as excluded instrument. A concern to the exclusion restriction when identifying peer effects through intransitive triads, discussed in Section 2.2, is that there may exist a direct channel between the characteristics of friends of friends and one's own actions (not only an indirect channel through the actions of one's friend). The degree to which this is a concern depends on the context (De Giorgi et al., 2010). In Section 5.3, we

Table 6: Effects of Distance to Disruptive Peers on Cognitive Scores at Ages 13, 16 and 19

	IQ Components at 13			*Grades at 16*	*Tests at 19*	
	Crystallized Int.		Fluid Int.			
	Verbal	Numeric	Spatial	Marks 9	Verbal	Spatial
Full Sample						
CWC in Class	-1.196**	-1.630**	0.458	-0.105		
	(0.551)	(0.706)	(0.692)	(0.088)		
Not Connected	0.995**	1.373**	-0.392	0.045		
	(0.502)	(0.665)	(0.638)	(0.089)		
Distance to Peer	0.401**	0.635**	-0.152	0.047		
	(0.185)	(0.264)	(0.250)	(0.036)		
Observations	7,032	7,032	7,032	6,751		
Males						
CWC in Class	-1.404*	-2.201*	0.200	-0.178	-0.141	-0.015
	(0.822)	(1.202)	(1.151)	(0.125)	(0.175)	(0.182)
Not Connected	1.153	1.895	0.374	0.147	0.101	0.040
	(0.785)	(1.223)	(1.079)	(0.120)	(0.166)	(0.170)
Distance to Peer	0.480*	0.938**	-0.007	0.095**	0.043	0.020
	(0.280)	(0.475)	(0.434)	(0.047)	(0.066)	(0.069)
Observations	3,450	3,450	3,450	3,309	3,156	3,164
Females						
CWC in Class	-0.798	-0.931	1.245	-0.201		
	(0.927)	(1.085)	(0.913)	(0.136)		
Not Connected	0.679	0.757	-1.629*	0.117		
	(0.929)	(1.031)	(0.913)	(0.134)		
Distance to Peer	0.242	0.365	-0.467	0.067		
	(0.328)	(0.416)	(0.339)	(0.051)		
Observations	3,582	3,582	3,582	3,442		

Note: Data from Stockholm Birth Cohort. Sample excludes classes with less than 7 students and schools with special education classrooms. *CWC in Class* takes on value 1 if there is a disruptive peer whose parents underwent an investigation for abuse and neglect by the child welfare committee (CWC) in the classroom and zero otherwise; *Not Connected* takes on value 1 if despite there being a disruptive peer in the classroom he/she does not belong to the social network of the student; and *Distance to Peer* stands for the geodesic distance between the student and the closest CWC peer. All regressions include school fixed effects and control for gender, 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. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

document evidence supporting the view that students are directly affected by their nearest peers and the drastic decrease in the effect with social distance is suggestive of friends' friend's characteristics affecting the student's actions only indirectly. As the distance results suggest that the effect of a disruptive peer has almost entirely dissipated once he/she is three or more edges away, the arguments for why the exclusion restriction should hold are stronger for the excluded instrument $\hat{\mathbf{G}}(\mathbf{Z})^3\mathbf{X}$. Therefore, we provide results for both models.

The findings suggest that there are statistically and economically significant contextual effects on both verbal and numeric ability of having disruptive peers in the classroom. The characteristics of a disruptive peer directly reduce the verbal ability score by an average of around 3 points. That represents roughly three quarters of the total peer-effect estimated in the reduced form. We find no endogenous effects that would snowball the negative disruptive peer characteristics through peers' verbal ability scores. As to numeric ability, we find by and large the same pattern. The negative contextual effect is in this case on average roughly -5 points. This is surprisingly close to the total peer-effect estimated in the reduced form (-4.747). Again, no endogenous peer effects are found on the student's numeric ability.[29] Taken together, the results on verbal and numeric test scores suggest that the contextual effect $\beta_{\bar{x}}$ accounts for almost all of the reduced-form coefficient $(I_r - \beta_{\bar{y}}\mathbf{J}_r)^{-1}\beta_{\bar{x}}$ of equation (3).

As one would expect, the model using fluid intelligence (i.e., spatial ability test score), as outcome, behaves differently from the models explaining verbal and numeric ability.

[29] One possible reason for our null findings on the endogenous effect could be the potential weakness of our instruments. As pointed out by Gibbons and Overman (2012), the 2SLS approach using intransitive triads has the potential problem of $\mathbf{G}$ and $\mathbf{G}^2$ being strongly correlated because friends tend to befriend friends of friends. However, our first stage results, presented in Table E.1 in the Appendix show that our instruments have substantial independent variation and are strong predictors of the endogeonus variable. This reflects the fact that collinearity between the instruments is less of a concern in small networks like classrooms, which have significant variation in the composition of friendship triads.

Table 7: Structural Estimation Results

	IQ Components at 13						*Grades at 16*	
	Crystallized Intelligence				Fluid Intelligence			
	Verbal		Numeric		Spatial		Marks 9	
Gy	-0.092	0.067	-0.474	-0.181	-1.254***	-1.081***	-4.163	1.499
	(0.153)	(0.165)	(0.302)	(0.310)	(0.370)	(0.380)	(5.641)	(3.420)
X	-1.733***	-1.680***	-2.175***	-2.070***	-1.056***	-1.043***	-43.782**	-28.337***
	(0.310)	(0.311)	(0.413)	(0.406)	(0.395)	(0.385)	(17.635)	(10.465)
$\hat{\mathbf{G}}(\mathbf{Z})\mathbf{X}$	-3.429***	-2.683**	-5.573***	-4.069*	-3.625**	-3.190*	-275.207	56.226
	(1.303)	(1.336)	(2.103)	(2.115)	(1.684)	(1.666)	(331.568)	(200.872)
Obs	7,481	7,481	7,493	7,493	7,481	7,481	7,153	7,153
Instruments for **Gy**								
$\hat{\mathbf{G}}(\mathbf{Z})^2\mathbf{X}$	✓		✓		✓		✓	
$\hat{\mathbf{G}}(\mathbf{Z})^3\mathbf{X}$	✓	✓	✓	✓	✓	✓	✓	✓

Note: Data from Stockholm Birth Cohort. Table present the estimates of structural model (13). The coefficients associated with **GY** represent the endogenous peer-effect. The coefficients associated with **GX** represent the exogenous peer-effects. $\hat{\mathbf{G}}(\mathbf{Z})$ is the predicted adjacency matrix obtained from (11). The notation in the table abstracts from the within-transformation **H** for clarity. Sample excludes classes with less than 7 students and schools with special education classrooms. Final sample is comprised by 387 classrooms in 118 schools. Estimations include school fixed-effects. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Both the contextual and the endogenous effects are negative. We do not have a direct interpretation for the results but note that we did not find a significant total peer-effect in the reduced form for spatial ability. Further, we do not find evidence for a significant peer effects for grades at age 16 (ninth grade).

Our results on crystalized intelligence show the importance of disentangling the endogenous effect from the contextual effect of having a disruptive peer in the classroom. The finding of contextual effect of having disruptive peers, isolated from the feedback effect opens up the lid of the black box of negative externalities of disruptive peers also documented in this study in Section 5.1. We find that most, if not all, of the effect disruptive peers have on their classmates comes directly from their disruptiveness, and not indirectly through their lower cognitive outcomes. The effect does not seem to propagate indirectly and hence does not create a social multiplier effect. In that sense, given that social links are the channels for the diffusion of the effect, our findings suggest that the cliqueness of social networks in classrooms prevents the propagation of negative externalities.

Understanding the nature of the peer effects at work has implications for designing optimal policy. The finding of negative contextual effects but no social multiplier suggests that disruptive peers directly affect achievement through affecting, e.g., the teacher's pedagogical work and consuming teachers' time and energy. However, the absence of an endogenous effect, and hence a social multiplier, leaves little scope for distribution of disruptive peers evenly across classrooms in order to minimize snowballing.

6 Robustness

In this section, we present a number of robustness checks and report results of tables which are found in Appendix D.

6.1 Model selection

A concern that one faces when applying the identification strategy proposed by Bramoullé et al. (2009) in the classroom context is that the exclusion restriction of the outcome equation (13) regarding intransitivity of triads may not hold in reality due to unobserved direct interaction between an individual and her friends' friends. We provide tentative evidence in Subsection 5.3 in support of the credibility of the exclusion restriction. In this subsection, we provide another indirect robustness check of the credibility of the intransitivity of triads by contesting our model selection against one in which all students in the classroom interact with each other equally, using a spatial J-test (Kelejian, 2008; Kelejian and Piras, 2011; Liu et al., 2014). Evidence in favor of our preferred model selection would support the understanding that not all students in the classroom interact with each other equally, and hence that the socio-matrix that fits the data best for classroom interactions with troubled peers among Swedish sixth graders is the network model with idiosyncratic friendship links.

The proposed alternative identification strategy solves the reflection problem in the linear-in-means model context based on group size variation and does hence not assume intransitive triads. Instead, Lee (2007) and Boucher et al. (2012) offer a solution in the simplest of settings: a cross-section of data with peers who interact in a group and are affected by common group factors. Identification stems from sufficient group size variation. Sufficient group size variation ensures $\mathbf{H}_s, \mathbf{H}_s\mathbf{J}_r$ and $\mathbf{H}_s\mathbf{J}_r^2$ will be linearly independent (Lee, 2007; Boucher et al., 2012). Insufficient classroom size variation, namely few distinct classroom sizes, manifests in multicollinearity. In line with Tatsi (2015), we subtract observations from their respective school mean instead of classroom. This both increases variation and leverages upon the idiosyncratic within-school across-classroom variation of disruptive peers.[30]

The results of the J-test of the model based on intransitive triads, in which we include the prediction from the model based group size variation, are presented in Table D.2. They suggest that the mechanism underlying the linear-in-means model based on average classroom peer effects is not relevant for the outcome of the model based on network interactions that identifies the parameters from intransitive triads.

6.2 Missing links

Even though all students in the classroom were asked to nominate their best friends the network is likely to suffer from missmeasurement. The number of links per student (i.e., edges per node) is going to be top coded due to the students only being allowed to nominate a maximum of three best friends. We observe that roughly 60 percent of the students nominated three best friends, suggesting that many would have nominated more had they had the chance to.

Censoring of links (edges) affects network characteristics such as the average number of direct friends (degree) and distance (path length) between students (nodes) (Kossinets, 2006). The concern when applying the identification strategy based on intransitive

[30]For a more detailed exposition of the method, see Appendix D and for the results of the structural model with the leave-out mean, see Table D.1.

triads to the linear-in-means model (Bramoullé et al., 2009; De Giorgi et al., 2010) is that some links are missing due to top coding and hence truly transitive triads appear falsely to be intransitive. Chandrasekhar and Lewis (2016) and Griffith (2017) 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. Both papers offer solutions that are not particularly tractable in our application. Instead in Table D.3 we turn to Patacchini et al. (2017) and conduct a variant of their robustness check by assigning direct links to friends of friends (i.e., closing observed intransitive triads into transitive ones). First, we add a friendship nomination to every student whose number of friends was top coded. We assign that link to the classmate with the closest $\mathbf{Z}$ vector who is currently not her friend. We then altered the procedure by adding one friendship nomination to every student who nominated three friends *and* has an estimated degree heterogeneity above the 50th (75th) percentile. Lastly, we include an additional friendship nomination to every student who nominated three friends by assigning the friendship nomination to a friend's friend. In all cases the added friendship link corresponds to the classmate with the highest predicted probability of friendship among those who are not reported as friends. The results remain remarkably robust to this test of bias due to missing links.

6.3 Exclusion restriction of the link formation regression

Even though the student allocation to classrooms solves the problem of correlated effects in the reduced-form model in Subsection 4.1, it does not do that when measuring the social interactions based on the observed social networks as we do in our distance and structural models in Subsections 4.3 and 4.4 respectively. Since individuals are likely to befriend each other based on characteristics that may affect the outcomes, our estimates of the structural model (13) may suffer from correlated effects. We dealt with this by replacing the actual links of the adjacency matrix by predicted links that come from a link formation regression.

In the link formation regression (11) the dyadic variables—whether student i and j are from the same gender, are first born or not, have a mother who received prenatal care by a physician, come from same parental SES at birth, are born into the same neighborhood—are defined at different aggregation level (i.e., for each i, j cell or dyad), as compared to outcome in the estimating equations (12) and (13) and are hence arguably exogenous predictors of link formation. The identifying assumption is thus that links are determined by a dyadic model whereas the outcomes are determined by a linear-in-means model 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, it may be the case that the aggregate differences in one's characteristics compared with those of one's friend directly influence the outcome. This does not directly invalidate our identification of the link formation but would intuitively be contradicting the exclusion restriction. We address this concern by running an informal test of the exclusion restriction proposed by Hsieh and Kippersluis (2018) that is in the spirit of Altonji et al. (2005) who test the stability of coefficients with respect to observable and unobservable characteristics. We include in the outcome equation (12) the share of children in the classroom who are similar to student i along the dyad-specific homophily variables variables, Z_{ij}, (i.e., same gender, neighborhood, father's occupational status at birth, etc.), $\sum_{j=1}^{n^r} \frac{Z_{ij}}{n^r-1}$. These control variables proxy for how similar to (or deviant from) the classmates the student is on average. A student that is deviant from the average classmate in terms of the dyadic homophily variables might not only find it difficult to befriend others but also score worse than average on achievement tests, which would go against the exclusion restriction of the dyadic homophily variables from the outcome equation. The results presented in Appendix D suggest remarkable stability of the estimates. The results reported in D.3 are very close to the benchmark estimates in Table 6.[31]

6.4 Placebo regressions

To further test whether it is our exogenous variation in exposure to disruptive peers that drives the results of effect heterogeneity reported in Table 6 we run our distance model (12) using as placebo outcomes number of household members in 1964, number of rooms in the apartment in 1960, postpartum stay (days) of the own mother after the cohort member's birth and own height. Not finding any significant evidence of distance to a disruptive peer affecting these outcomes negatively would be suggestive of our main results not being produced by alternative processes related to selection or classroom manipulation by parents. Table D.5 of Appendix D.4 shows no evidence of distance to the disruptive peers having an effect on the proposed placebo outcomes.

7 Conclusions

This study explores the effects of disruptive peers in primary school on cognitive outcomes. Using the standard reduced-form approach to identify the total peer-effect from

[31] Of the 9 "average distance to classmates" variables, only the coefficient of average deviance from classmates in terms of fathers socioeconomic status at birth has a significant t-value at 5 percent level in the models reported in D.3. The others are not significant at 15 percent level in any of the models. The coefficients and standard errors are available from the authors.

within-school variation in the fraction of disruptive peers, we find that disruptive peers will negatively affect cognitive short-term outcomes such as verbal and numeric ability in sixth grade of primary school as well as the grade point average in ninth rade of high school. When separating between female and male disruptive peers we see that disruptive boys exert negative externalities on classmates of both genders while disruptive girls only seem to confer negative spillovers on their female classmates.

We move beyond the reduced form of the linear-in-means model in an attempt to explore the mechanisms through which the disruptive peers impact their classmates. We use sociometric data to show that social connections matter in the transmission of the peer-effect. Classmates who are directly connected to the disruptive peer face substantially stronger negative effects on cognitive outcomes than those more remotely, or not at all, connected to the disruptive peer. Our results show that being three friends (as measured by number of edges) away from the disruptive peer is almost equivalent to not having him or her at all in the classroom. These finding suggest that disruptive peers do not impede everyone's learning and cognitive development equally. Instead, their troubled experiences involving abuse and neglect, their externalizing behavior, substance abuse and scholastic motivation confer negative spillovers primarily on their closest friends and but these spillovers dissipate quickly as the social distance to the peer increase.

Furthermore, we estimate the structural parameters of the linear-in-means model using methods that help us overcome the reflection problem and provide a way to disentangle the endogenous effects of peer's actions from the contextual effect of peer characteristics. We find that the direct contextual effect of disruptive peer characteristics accounts for three quarters to as good as all of the total peer-effect. Thus, we find no evidence of a social multiplier generated by the endogenous effect.

To the best of our knowledge this study is the first study to identify the endogenous and contextual peer effects of having disruptive peers on cognitive outcomes. Importantly, our estimation strategy deals with endogenous network formation. We extended the reduced-form model to account for effect heterogeneity depending on the distance to the disruptive peer and show that the negative peer effect fades out quickly the less connected the students are to this peer. Our results of the structural model show that the total peer-effect documented in the reduced form is mainly due to contextual effects of peer characteristics, not their performance. Our main conclusion is that the disruptive students' characteristics confer significant contextual peer effects which account for almost all of the negative total peer-effect. We do not find evidence of the treatment initiating a snowball effect that would propagate throughout the network through the actions of peers and the feed back loop of actions.

Understanding the nature of the peer effects at work has implications for designing optimal policy. The finding of negative contextual effects but no social multiplier suggests that disruptive peers directly affect achievement through affecting the teacher's pedagogical work and consuming the teacher's time and energy. However, the absence of an endogenous effect, and hence a social multiplier, leaves little scope for policies intended to distribute disruptive peers evenly across classrooms in order to minimize snowballing. Increasing teaching resources, which would allow a greater focus on challenging students, might be one way to contain the direct contextual harms of disruptive peers. Another fairly evident policy prescription in order to reduce the negative contextual effects is to try to change the context, in this case disruptiveness. Targeting children exposed to maltreatment at an early age and encouraging compensating investments will generate positive spillover effects that by far exceed the direct remediating effects on these children's own outcomes.

References

Aizer, A. (2008). Peer Effects and Human Capital Accumulation: the Externalities of ADD. 1, 4

Aizer, A. (2011). Poverty, Violence, and Health: The Impact of Domestic Violence During Pregnancy on Newborn Health. *Journal of Human Resources*, 46(3):518–538. 3

Almlund, M., Duckworth, A. L., Heckman, J., and Kautz, T. (2011). Personality Psychology and Economics. *Handbook of the Economics of Education*, 4:1–181. 3.1

Altonji, J. G., Elder, T. E., and Taber, C. R. (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy*, 113(1):151–184. 6.3

Ammermueller, A. and Pischke, J. S. (2009). Peer Effects in European Primary Schools: Evidence from the Progress in International Reading Literacy Study. *Journal of Labor Economics*, 27(3):315–348. 4

Angrist, J. (2014). The perils of peer effects. *Labour Economics*, 30:98–108. 4.1

Attanasio, O., Barr, A., Cardenas, J. C., Genicot, G., and Meghir, C. (2012). Risk Pooling, Risk Preferences, and Social Networks. *American Economic Journal: Applied Economics*, 4(2):134–167. 2.1

Auerbach, E. (2019). Identification and Estimation of a Partially Linear Regression Model using Network Data. *arXiv.org*, page arXiv:1903.09679. 13

Badev, A. (2018). Nash Equilibria on (Un)Stable Networks. *arXiv.org*, page arXiv:1901.00373. 10, 4.2, 24

Banerjee, A., Chandrasekhar, A. G., Duflo, E., and Jackson, M. O. (2013). The Diffusion of Microfinance. *Science*, 341(6144):1236498–1236498. 2

Bifulco, R., Fletcher, J. M., and Ross, S. L. (2011). The effect of classmate characteristics on post-secondary outcomes: Evidence from the Add Health. *American Economic Journal: Economic Policy*, 3(1):25–53. 4

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2013). Under Pressure? The Effect of Peers on Outcomes of Young Adults. *Journal of Labor Economics*, 31(1):119–153. 4

Blume, L., Brock, W., Durlauf, S., and Ioannides, Y. (2011). Identification of social interactions. *Handbook of Social Economics*, pages 854–965. 2.1, 7, 8, 17

Blume, L. E., Brock, W. A., Durlauf, S. N., and Jayaraman, R. (2015). Linear Social Interactions Models. *Journal of Political Economy*, 123(2):444–496. 7, 13

Borghans, L., Duckworth, A. L., Heckman, J. J., and Ter Weel, B. (2008). The Economics and Psychology of Personality Traits. *Journal of Human Resources*, 43(4):972–1059. 3.1

Boucher, V., Bramoullé, Y., Djebbari, H., and Fortin, B. (2012). Do Peers Affect Student Achievement? Evidence From Canada Using Group Size Variation. *Journal of applied econometrics*, 29(1):91–109. 6.1, D.1.1

Bramoullé, Y., Djebbari, H., and Fortin, B. (2009). Identification of peer effects through social networks. *Journal of Econometrics*, 150(1):41–55. 1, 2.1, 2.1, 2.1, 2.2, 2.2, 4.4, 4.4, 26, 6.1, 6.2, D.1

Brock, W. and Durlauf, S. (2001). Discrete choice with social interactions. *The Review of Economic Studies*, 68(2):235–260. 2.1

Calvó-Armengol, A., Patacchini, E., and Zenou, Y. (2009). Peer Effects and Social Networks in Education. *The Review of Economic Studies*, 76(4):1239–1267. 2, 7

Carlson, B. E. (2000). Children Exposed to Intimate Partner Violence Research Findings and Implications for Intervention. *Trauma, Violence, & Abuse*, 1(4):321–342. 3

Carrell, S. E. and Hoekstra, M. L. (2010). Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone's Kids. *American Economic Journal: Applied Economics*, 2(1):211–228. 1, 3, 4, 2.1, 4.1, 5.1

Carrell, S. E., Hoekstra, M. L., and Kuka, E. (2018). The Long-Run Effects of Disruptive Peers. *American Economic Review*, 108:3377–3415. 1, 4, 2.1

Carrell, S. E., Sacerdote, B. I., and West, J. E. (2013). From natural variation to optimal policy? The importance of endogenous peer group formation. *Econometrica*, 81(3):855–882. 1, 8, 2.1

Chandrasekhar, A. and Lewis, R. (2016). Econometrics of sampled networks. 6.2

Charbonneau, K. B. (2017). Multiple fixed effects in binary response panel data models. *The Econometrics Journal*, 20(3):S1–S13. 4.2

Coleman, J. S. (1964). *Introduction to mathematical sociology*. Free Press. 2.1

Comola, M. and Prina, S. (2015). Treatment Effect Accounting for Network Changes: Evidence from a Randomized Intervention. 5

Currie, J. and Widom, C. S. (2010). Long-Term Consequences of Child Abuse and Neglect on Adult Economic Well-Being. *Child maltreatment*, 15(2):111–120. 3

Dahl, G. B., Løken, K. V., and Mogstad, M. (2014). Peer Effects in Program Participation. *American Economic Review*, 104(7):2049–2074. 8

De Giorgi, G., Pellizzari, M., and Redaelli, S. (2010). Identification of Social Interactions through Partially Overlapping Peer Groups. *American Economic Journal: Applied Economics*, 2(2):241–275. 1, 6, 2.2, 5.4, 6.2

de Paula, Á. (2017). Econometrics of Network Models. In Honore, B., Pakes, A., Piazzesi, M., and Samuelson, L., editors, *Advances in Economics and Econometrics*, pages 268–323. Cambridge University Press, Cambridge. 1, 5, 2.1, 24

Doyle Jr, J. J. (2007). Child Protection and Child Outcomes: Measuring the Effects of Foster Care. *American Economic Review*, 97(5):1583–1610. 3

Doyle Jr, J. J. (2008). Child Protection and Adult Crime: Using Investigator Assignment to Estimate Causal Effects of Foster Care. *Journal of Political Economy*, 116(4):746–770. 3

Eriksen, T. L. M., Nielsen, H. S., and Simonsen, M. (2014). Bullying in Elementary School. *Journal of Human Resources*, 49(4):839–871. 3

Figlio, D. (2005). Boys named Sue: disruptive children and their peers. 4

Fredriksson, P., Ockert, B., and Oosterbeek, H. (2012). Long-Term Effects of Class Size. *The Quarterly Journal of Economics*, 128(1):249–285. 21

Gibbons, S. and Overman, H. G. (2012). Mostly Pointless Spatial Econometrics? *Journal of Regional Science*, 52(2):172–191. 29

Goeree, J. K., McConnell, M. A., Mitchell, T., Tromp, T., and Yariv, L. (2010). The 1/ dLaw of Giving. *American Economic Journal: Microeconomics*, 2(1):183–203. 28

Goldsmith-Pinkham, P. and Imbens, G. W. (2013). Social Networks and the Identification of Peer Effects. *Journal of Business & Economic Statistics*, 31(3):253–264. 1, 2.1, 2.1, 13, 4.2

Gould, E. D., Lavy, V., and Daniele Paserman, M. (2009). Does Immigration Affect the Long-Term Educational Outcomes of Natives? Quasi-Experimental Evidence. *The Economic Journal*, 119(540):1243–1269. 4

Graham, B. S. (2015). Methods of Identification in Social Networks. *Annual Review of Economics*, 7(1):465–485. 10, 4.2, 24

Graham, B. S. (2017). An Econometric Model of Network Formation With Degree Heterogeneity. *Econometrica*, 85(4):1033–1063. 2.1, 10, 2.1, 4.2, 24

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

Griffith, A. (2017). How Many Friends Do You Have? An Empirical Investigation into Censoring-Induced Bias in Social Network Data. 6.2

Guryan, J., Kroft, K., and Notowidigdo, M. J. (2009). Peer Effects in the Workplace: Evidence from Random Groupings in Professional Golf Tournaments. *American Economic Journal: Applied Economics*, 1(4):34–68. 4.1, 18, 2

Heckman, J. (1979). Sample specification bias as a selection error. *Econometrica*, 47(1):153–161. 13

Heckman, J. J. and Robb, R. (1986). Alternative Methods for Solving the Problem of Selection Bias in Evaluating the Impact of Treatments on Outcomes. In *Drawing Inferences from Self-Selected Samples*, pages 63–107. Springer, New York, NY, New York, NY. 13

Holt, S., Buckley, H., and Whelan, S. (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. 3

Hoxby, C. (2000). Peer Effects in the Classroom: Learning from Race and Gender. 1, 4

Hsieh, C.-S. and Kippersluis, H. (2018). Smoking initiation: Peers and personality. *Quantitative Economics*, 9(2):825–863–863. 5, 6.3

Hsieh, C.-S. and Lee, L.-f. (2015). A Social Interactions Model with Endogenous Friendship Formation and Selectivity. *Journal of applied econometrics*, 31(2):301–319. 1, 2.1, 13, 4.2

Jackson, M. O. (2010). *Social and Economic Networks*. Princeton University Press. 2.1, 11, 4.2

Jackson, M. O. and Rogers, B. W. (2007). Meeting Strangers and Friends of Friends: How Random Are Social Networks? *American Economic Review*, 97(3):890–915. 2.1, 9, 2.3

Jackson, M. O. and Wolinsky, A. (1996). A Strategic Model of Social and Economic Networks. *Journal of Economic Theory*, 71(1):44–74. 10

Jackson, M. O. and Yariv, L. (2011). Diffusion, Strategic Interaction, and Social Structure. Elsevier. 1, 2

Jackson, M. O. and Zenou, Y. (2015). Games on Networks. Elsevier. 7

Johnsson, I. and Moon, H. R. (2017). Estimation of Peer Effects in Endogenous Social Networks: Control Function Approach. 13

Kelejian, H. and Prucha, I. (1998). A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances. *Journal of Real Estate Finance and Economics*, 17(1):99–121. 2.2, 26

Kelejian, H. H. (2008). A spatial J-test for model specification against a single or a set of non-nested alternatives. *Letters in Spatial and Resource Sciences*, 1(1):3–11. 6.1, D.1

Kelejian, H. H. and Piras, G. (2011). An extension of Kelejian's J-test for non-nested spatial models. *Regional Science and Urban Economics*, 41(3):281–292. 6.1, D.1

Kelejian, H. H. and Prucha, I. R. (1997). Estimation of Spatial Regression Models with Autoregressive Errors by Two-Stage Least Squares Procedures: A Serious Problem. *International Regional Science Review*, 20(1-2):103–111. 26

König, M. D., Liu, X., and Zenou, Y. (2018). R&D Networks: Theory, Empirics and Policy Implications. *The Review of Economics and Statistics*, 0(0). 5, 2.1, 2.2, 4.3, 4.4

Kossinets, G. (2006). Effects of missing data in social networks. *Social Networks*, 28(3):247–268. 6.2

Lavy, V., Paserman, M. D., and Schlosser, A. (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. 4, 2.1, 8

Lee, L.-f. (2007). Identification and estimation of econometric models with group interactions, contextual factors and fixed effects. *Journal of Econometrics*, 140(2):333–374. 6.1, D.1.1, D.1.1

Leider, S., Möbius, M. M., Rosenblat, T., and Do, Q.-A. (2009). Directed altruism and enforced reciprocity in social networks. *The Quarterly Journal of Economics*, 124(4):1815–1851. 28

Lin, X. (2010). Identifying Peer Effects in Student Academic Achievement by Spatial Autoregressive Models with Group Unobservables. *Journal of Labor Economics*, 28(4):825–860. 6, 2.1, 4.2

List, J., Momeni, F., and Zenou, Y. (2018). Are Measures of Early Education Programs Too Pessimistic? Evidence from a Large-Scale Field Experiment. 2

Liu, X., Patacchini, E., and Zenou, Y. (2014). Endogenous peer effects: local aggregate or local average? *Journal of Economic Behavior & Organization*, 103:39–59. 4.2, 6.1, D.1

Manski, C. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, 60(3):531–542. 1, 2, 2.1, 2.1, D.1.1

Manski, C. F. (2013). Comment. *Journal of Business & Economic Statistics*, 31(3):273–275. 1, 2.1

McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, 27:415–444. 2.1, 4.2

Moylan, C. A., Herrenkohl, T. I., Sousa, C., Tajima, E. A., Herrenkohl, R. C., and Russo, M. J. (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. 3

Opper, I. M. (2019). Does Helping John Help Sue? Evidence of Spillovers in Education. *American Economic Review*, 109(3):1080–1115. 2

Patacchini, E., Rainone, E., and Zenou, Y. (2017). Heterogeneous peer effects in education. *Journal of Economic Behavior & Organization*, 134:190–227. 6, 2.1, 4.2, 6.2

Patacchini, E. and Zenou, Y. (2012). Juvenile Delinquency and Conformism. *Journal of Law, Economics, and Organization*, 28(1):1–31. 7

Sacerdote, B. (2011). Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far? Elsevier. 1, 1, 8

Sarzosa, M. and Urzua, S. (2015). Bullying among Adolescents: The Role of Cognitive and Non-Cognitive Skills. *NBER Working Paper*, 21631. 3

Schennach, S. M. (2018). Long Memory via Networking. *Econometrica*, 86:2221–2248. 2

Sheng, S. (2016). A Structural Econometric Analysis of Network Formation Games. *MIMEO. UCLA*, pages 1–70. 10, 4.2, 24

Stenberg, S. A. and Vagerö, D. (2006). Cohort Profile: The Stockholm birth cohort of 1953. *International Journal of Epidemiology*, 35(3):546–548. 14

Tatsi, E. (2015). Endogenous Social Interactions: Which Peers Matter? *Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung - Theorie und Politik - Session: Understanding the nature of peer effects*, 1. 6.1

Wolfe, D. A., Crooks, C. V., Lee, V., McIntyre-Smith, A., and Jaffe, P. G. (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. 3

Appendix

A Abuse and neglect and own outcomes

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	Appl./Accept.	Any Post
	Verbal	Spatial	Numeric	Education	Upper Secondary	Secondary
Abused & Neglected	-2.361***	-1.468***	-2.498***	-0.711***	-0.205***	-0.123***
	(0.243)	(0.267)	(0.301)	(0.094)	(0.018)	(0.019)
Observations	11,434	11,431	11,444	11,501	12,047	11,501
Acting out during adolescence						
	Maladjustment	Drugs	Violence	Vandalism	Drunk	Delinquency
Abused & Neglected	0.084***	0.063***	0.035***	0.011***	0.049***	0.123***
	(0.007)	(0.006)	(0.006)	(0.004)	(0.007)	(0.008)
Observations	12,047	12,047	12,047	12,047	12,047	12,047

Note: Data from Stockholm Birth Cohort. The regression includes school fixed effects and gender controls. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

B Balancing test with binary exposure

Table B.1: Exposure to Troubled Peers and Own Characteristics

	Gender	SocAid Recipt	Birth Weight	Mother Age	Owner Dwell	Dwell Size	Older Sibls
At least 1 peer with CWC evaluation							
All Peers	-0.006	-0.005	0.000	-0.001	0.000	-0.018	0.000
	(0.008)	(0.016)	(0.001)	(0.001)	(0.025)	(0.032)	(0.005)
Male Peers	0.002	0.003	0.000	0.000	0.001	-0.032	-0.009
	(0.010)	(0.016)	(0.001)	(0.001)	(0.025)	(0.036)	(0.005)
Female Peers	-0.010	-0.012	-0.001	0.000	0.014	-0.014	0.006
	(0.007)	(0.018)	(0.001)	(0.001)	(0.022)	(0.025)	(0.005)
Observations	7,459	7,459	6,322	6,486	7,459	7,459	7,459

Note: Each cell reports a separate regression. All regressions include school fixed effects, exclude classes with < 7 students and schools with special education classrooms. Sample restricted to individuals who $CWC = 0$. Standard errors in parentheses clustered at the school level. *** $p < 0.001$, ** $p < 0.05$, * $p < 0.1$.

C Link formation: Homophily

Table C.1: Summary Statistics of the Homophily Variables

Variable	**Mean**	**Std. Dev.**	**N**
Female	0.509	0.500	7,826
Parents Social Aid Recipients	0.121	0.326	7,826
Own Dwelling	0.190	0.392	7,826
Older Siblings Deceased	0.158	0.366	6,651
First Pregnancy	0.386	0.487	6,651
Prenatal Care by Doctor	0.423	0.494	6,638
Natal Health 1	-0.021	1.398	6,431
Natal Health 2	-0.004	1.300	6,431
Father's Occupational Status in 1953			
Upper & Upper Middle Class	0.146	0.353	7,683
Middle Class	0.336	0.473	7,683
Lower Middle Class	0.065	0.249	7,683
Skilled Workers	0.280	0.449	7,683
Unskilled Workers	0.172	0.378	7,683

Note: Data from Stockholm Birth Cohort. Sample excludes classes with less than 7 students and schools with special education classrooms. *Natal Health 1* and *Natal Health 2* are two indices that summarize the prenatal and perinatal health of the child as evidenced by the birth records in 1953. They are the first two principal components of a system of measures that include: age of mother at the time of giving birth, whether delivery occurred in a hospital, length of postpartum stay in hospital, whether mother suffered from pre-eclampsia during pregnancy, whether diagnosed with other health conditions during pregnancy, whether mother was diagnosed with anemia during pregnancy, whether mother had fever during delivery, whether the child was born through C-section, whether the delivery was facilitated by forceps, birthweight, and length at birth. *Natal Health 1* and *Natal Health 2* collect 14% and 12.3% of the measurement system variation respectively.

D Robustness

D.1 Model selection

A concern that one faces when applying the identification strategy proposed by Bramoullé et al. (2009) in the classroom context is that the exclusion restriction regarding intransitivity of triads may not hold in reality due to unobserved direct interaction between an individual and her friends' friends. We provide tentative evidence in Subsection (5.3) in support of the credibility of the exclusion restriction. In this subsection, we provide another indirect robustness check of the credibility of the intransitivity of triads by contesting our model selection against one in which all students in the classroom interact with each other equally much using a spatial J-test (Kelejian, 2008; Kelejian and Piras, 2011; Liu et al., 2014). Evidence in favor of our preferred model selection would support the understanding that not all students in the classroom interact with each other equally much and hence that the socio-matrix that fits the data best for classroom interactions with disruptive peers among Swedish sixth graders is the network model with idiosyncratic friendship links. Let us begin by outlining our identification strategy of peer effects in a context where all students in the classroom interact with each other equally much.

D.1.1 Identification using group size variation

Even in settings where all members of the reference group are linked to one another, such as in (2) where the socio-matrix was of the form:

$$\mathbf{J}_r = (n_r - 1)^{-1}(\iota_{n_r}\iota'_{n_r} - I_{n_r}), \tag{14}$$

the parameters of the linear in means model can be identified given sufficient group size variation (Lee, 2007; Boucher et al., 2012). Under the generic reference group structure of (14) where all individuals are linked to one another, we have

$$\mathbf{J}_r^2 = \frac{1}{n_r - 1} I_{n_r} + \frac{n_r - 2}{n_r - 1} J_r. \tag{15}$$

If reference groups vary in size, then I_{n_r}, $\mathbf{J}_r$, $\mathbf{J}_r^2$ and $\mathbf{J}_r^3$ will be linearly independent (cf. Lee (2007)). With group fixed effects, identification requires three different group sizes, e.g., classrooms with 15, 18 or 20 students. Manski (1993) considered large groups with $n_r \to \infty$, in which case it is clear from (15) that $\mathbf{J}_r = \mathbf{J}_r^2$ and identification fails. The larger the groups are, the weaker is the identification. The structural equation of (2) can be estimated by instrumental variable methods in a similar way as (13). First, we redefine the sociomatrix to be $N \times N$ block-diagonal $\mathbf{J} = \mathrm{Diag}(\mathbf{J}_1, ..., \mathbf{J}_R)$ and then concatenate the estimating equation over all groups

$$\mathbf{Y} = \beta\mathbf{J}\mathbf{y} + \mathbf{X}\gamma + \mathbf{J}\mathbf{X}\delta + \eta, \tag{16}$$

where $\mathbf{Y}$ and $\mathbf{X}$ are stacked vectors of Y_r and X_r for $r = 1, ..., R$. Given sufficient group size variation and exclusive means, $E\left[\mathbf{J}\mathbf{Y}|X,\, J\right]$ is not perfectly collinear to $(\mathbf{X},\, \mathbf{J}\mathbf{X})$ and the model is identified and $\mathbf{J}^2\mathbf{X}$ can be used as a matrix of valid instruments for $\mathbf{J}\mathbf{X}$. We incorporate a block-diagonal matrix $\mathbf{H} = \mathrm{Diag}(\mathbf{H}_1, ..., \mathbf{H}_R)$ of the school within-transformation $\mathbf{H}_s = (I_{n_s} - \frac{1}{n_s}\iota_{n_s}\iota'_{n_s})$ as previously to control for school fixed effects.

Table D.1: Structural Results Using Leave-One-Out Mean

	IQ Components at 13			*Grades at 16*
	Crystallized Int		Fluid Int.	
	Verbal	Numeric	Spatial	Marks 9
Jy	0.670*	0.580	0.465	0.006***
	(0.361)	(0.501)	(0.540)	(0.002)
X	-1.624***	-1.908***	-0.961***	-31.924***
	(0.318)	(0.409)	(0.351)	(4.068)
JX	-0.477	-0.946	-0.269	-16.306
	(2.472)	(3.735)	(2.478)	(21.282)
Observations	7,481	7,493	7,481	7,153

Note: Data from Stockholm Birth Cohort. Table present the estimates of structural model (16)); that is using the leave-out mean as the model for peer-effects. The coefficients associated with **Jy** represent the endogenous peer-effect. The coefficients associated with **JX** represent the exogenous peer-efects. Endogenous variables **Jy** is instrumented using $\mathbf{J}^2\mathbf{X}$ and $\mathbf{J}^3\mathbf{X}$. The notation in the table abstracts from the within-transformation **H** for clarity. Sample excludes classes with less than 7 students and schools with special education classrooms. Final sample comprises 387 classrooms in 118 schools. Estimations include school fixed-effects. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table D.2: Robustness Check: J-Test Structural Estimation Results

	IQ Components at 13						*Grades at 16*	
	Crystallized Intelligence				Fluid Intelligence			
	Verbal		Numeric		Spatial		Marks 9	
$\mathbf{Gy}$	0.067	0.149	-0.181	-0.568	-1.081***	-2.061***	1.499	-1.564
	(0.165)	(0.312)	(0.310)	(0.759)	(0.380)	(0.777)	(3.420)	(4.883)
$\mathbf{X}$	-1.680***	-2.081***	-2.070***	0.025	-1.043***	3.118**	-28.337***	109.962
	(0.311)	(0.739)	(0.406)	(2.337)	(0.385)	(1.580)	(10.465)	(388.18)
$\hat{\mathbf{G}}(\mathbf{Z})\mathbf{X}$	-2.683**	-2.699**	-4.069*	-3.786**	-3.190*	-2.241	56.226	-55.140
	(1.336)	(1.321)	(2.115)	(1.916)	(1.666)	(1.630)	(200.87)	(101.41)
$\hat{\mathbf{Y}}(\mathbf{J}_r\mathbf{X})$		-0.253		1.121		4.407***		4.544
		(0.460)		(1.292)		(1.648)		(12.440)
Observations	7,481	7,481	7,493	7,493	7,481	7,481	7,153	7,153

Note: Data from Stockholm Birth Cohort. Table present the estimates of structural model (13). The coefficients associated with $\mathbf{Gy}$ represent the endogenous peer-effect. The coefficients associated with $\mathbf{GX}$ represent the exogenous peer-efects. Endogenous variables $\mathbf{Gy}$ is instrumented using $\hat{\mathbf{G}}^2\mathbf{X}(\mathbf{Z})$ and $\hat{\mathbf{G}}^3\mathbf{X}(\mathbf{Z})$, where $\hat{\mathbf{G}}(\mathbf{Z})$ is the predicted adjacency matrix obtained from (11). $\hat{\mathbf{Y}}(\mathbf{J}_r\mathbf{X})$ collects predicted outcome obtained from running model (16); that is using the the leave-one-out mean as the model for peer-effects. The notation in the table abstracts from the within-transformation $\mathbf{H}$ for clarity. Sample excludes classes with less than 7 students and schools with special education classrooms. Final sample comprises 387 classrooms in 118 schools. Estimations include school fixed-effects. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

D.2 Stability of the coefficients in the distance equation.

Table D.3: Testing Stability of the Coefficients in the Model on Effects of Distance to Disruptive Peers on Cognitive Scores

	IQ Components at 13				*Grades at 16*	*Tests at 19*	
	Crystallized Int.			Fluid Int.			
	Verbal	Numeric		Spatial	Marks 9	Verbal	Spatial
Full Sample							
CWC in Class	-1.191**	-1.629**		0.342	-0.106		
	(0.559)	(0.741)		(0.711)	(0.084)		
Not Connected	1.025**	1.416**		-0.227	0.053		
	(0.510)	(0.705)		(0.660)	(0.086)		
Distance to Peer	0.421**	0.660**		-0.078	0.054		
	(0.197)	(0.288)		(0.264)	(0.034)		
Observations	7,032	7,032		7,032	6,751		
Males							
CWC in Class	-1.410*	-2.511**		-0.174	-0.172	-0.162	-0.127
	(0.846)	(1.157)		(1.245)	(0.123)	(0.170)	(0.214)
Not Connected	1.207	2.309*		0.811	0.161	0.145	0.155
	(0.805)	(1.190)		(1.155)	(0.118)	(0.159)	(0.197)
Distance to Peer	0.519*	1.132**		0.191	0.106**	0.065	0.069
	(0.300)	(0.479)		(0.463)	(0.045)	(0.063)	(0.080)
Observations	3,450	3,450	3,450	3,309	3,156	3,164	
Females							
CWC in Class	-0.838	-0.485		1.418	-0.222		
	(0.988)	(1.106)		(0.891)	(0.137)		
Not Connected	0.766	0.392		-1.745**	0.140		
	(0.989)	(1.028)		(0.882)	(0.143)		
Distance to Peer	0.264	0.195		-0.512	0.073		
	(0.359)	(0.411)		(0.332)	(0.054)		
Observations	3,582	3,582		3,582	3,442		

Note: Data from Stockholm Birth Cohort. Sample excludes classes with less than 7 students and schools with special education classrooms. *CWC in Class* takes on value 1 if there is a disruptive peer whose parents underwent an investigation for abuse and neglect by the child welfare committee (CWC) in the classroom and zero otherwise; *Not Connected* takes on value 1 if despite there being a disruptive peer in the classroom he/she does not belong to the social network of the student; and *Distance to Peer* stands for the geodesic distance between the student and the closest CWC peer. All regressions include school fixed effects and control for the same sociodemographic characteristics as the benchmark distance estimations reported in Table 6 and the share of children in the classroom who are similar to student i in terms of the of the dyad-specific homophily variables Z_{ij} of the link formation estimating equation (11), reported in Table 5. Standard errors, shown in parentheses, are clustered at the school level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

D.3 Truncation in the Nomination of Friends

Table D.4: Structural Estimation Results: Adding Friendship Nominations

	IQ Components at 13 (Crystallized Intelligence)							
	Verbal				Numeric			
	All	Top 50	Top 25	Fr of Fr	All	Top 50	Top 25	Fr of Fr
$\mathbf{Gy}$	0.020	-0.028	-0.037	-0.092	-0.330	-0.400	-0.408	-0.511
	(0.193)	(0.179)	(0.174)	(0.157)	(0.371)	(0.349)	(0.340)	(0.314)
$\mathbf{X}$	-1.623***	-1.656***	-1.663***	-1.730***	-2.172***	-2.207***	-2.211***	-2.191***
	(0.319)	(0.317)	(0.316)	(0.311)	(0.438)	(0.431)	(0.428)	(0.417)
$\hat{\mathbf{G}}(\mathbf{Z})\,\mathbf{X}$	-3.197**	-3.510***	-3.548***	-3.601***	-5.824***	-6.156***	-6.144***	-5.988***
	(1.313)	(1.326)	(1.330)	(1.330)	(2.135)	(2.144)	(2.145)	(2.147)
Observations	7,407	7,406	7,405	7,481	7,420	7,419	7,418	7,493

Note: Data from Stockholm Birth Cohort. Table present the estimates of structural model (13) with the inclusion of friendship nominations additional to the ones reported in the data. In Columns labeled *All*, we include an additional friendship nomination to every student who nominated three (i.e., the maximum) friends. In Columns labeled *Top 50* (*Top 25*), we include an additional friendship nomination to every student who nominated three friends *and* has an estimated degree heterogeneity above the 50th (75th) percentile. In Columns labeled *Fr of Fr*, we also include an additional friendship nomination to every student who nominated three friends. But in this case, the friendship nomination given is a friend's friend. 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 $\mathbf{GY}$ represent the endogenous peer-effect. The coefficients associated with $\mathbf{GX}$ represent the exogenous peer-effects. $\hat{\mathbf{G}}(\mathbf{Z})$ is the predicted adjacency matrix obtained from (11). We use $\hat{\mathbf{G}}(\mathbf{Z})^{\mathbf{2}}\mathbf{X}$ and $\hat{\mathbf{G}}(\mathbf{Z})^{\mathbf{3}}\mathbf{X}$ as instruments for $\mathbf{GY}$. The notation in the table abstracts from the within-transformation $\mathbf{H}$ for clarity. Sample excludes classes with less than 7 students and schools with special education classrooms. Final sample is comprised by 387 classrooms in 118 schools. Estimations include school fixed-effects. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

D.4 Placebo regressions

Table D.5: Placebo test of the Model on Effects of Distance to Disruptive Peers on Cognitive Scores

	Number of individuals in the household (1964)	Number of rooms in the apartment (1960)	Height	Own mother's length of postpartum stay
Full Sample				
CWC in Class	0.008	0.012	0.025	0.022
	(0.087)	(0.093)	(0.238)	(0.115)
Not Connected	-0.011	-0.020	0.170	-0.043
	(0.087)	(0.078)	(0.238)	(0.113)
Distance to Peer	-0.019	-0.023	0.065	-0.005
	(0.034)	(0.032)	(0.095)	(0.044)
Observations	7,032	7,032	7,032	5,837
Males				
CWC in Class	-0.095	0.098	-0.155	0.079
	(0.137)	(0.145)	(0.364)	(0.160)
Not Connected	0.094	-0.074	0.493	-0.066
	(0.130)	(0.127)	(0.352)	(0.155)
Distance to Peer	0.004	-0.063	0.188	-0.006
	(0.053)	(0.050)	(0.148)	(0.058)
Observations	3,450	3,450	3,450	2,855
Females				
CWC in Class	0.118	0.020	0.422	-0.193
	(0.144)	(0.122)	(0.356)	(0.171)
Not Connected	-0.126	-0.067	-0.380	0.147
	(0.149)	(0.123)	(0.389)	(0.168)
Distance to Peer	-0.045	-0.020	-0.122	0.063
	(0.055)	(0.046)	(0.143)	(0.070)
Observations	3,582	3,582	3,582	2,982

Note: Data from Stockholm Birth Cohort. Sample excludes classes with less than 7 students and schools with special education classrooms. *CWC in Class* takes on value 1 if there is a disruptive peer whose parents underwent an investigation for abuse and neglect by the child welfare committee (CWC) in the classroom and zero otherwise; *Not Connected* takes on value 1 if despite there being a disruptive peer in the classroom he/she does not belong to the social network of the student; and *Distance to Peer* stands for the geodesic distance between the student and the closest CWC peer. All regressions include school fixed effects and control for the same sociodemographic characteristics as the benchmark distance estimations reported in Table 6 and the sum of the dyad-specific variables Z_{ij} controlled for in the link formation estimating equation (11), reported in Table 5, summed over all the student's friends. Standard errors, shown in parentheses, are clustered at the school level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

E Structural results: First stage

Table E.1: Structural Estimation: First Stage Results

	Gy *of IQ Components at 13*						**Gy** *of Grades at 16*	
	Crystallized Intelligence				Fluid Intelligence			
	Verbal		Numeric		Spatial		Marks 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)
Obs.	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. Table present the first stage estimates of structural results shown in Table 7. The coefficients associated with $\mathbf{GX}$ represent the exogenous peer-effects. $\hat{\mathbf{G}}(\mathbf{Z})$ is the predicted adjacency matrix obtained from (11). Sample excludes classes with less than 7 students and schools with special education classrooms.Estimations include school fixed-effects. Constant not shown. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.