Heterogeneity in Network Peer Effects*

Livia Shkoza[a]
University of Konstanz
GSDS

Derya Uysal[b]
University of Munich

Winfried Pohlmeier[c]
University of Konstanz
CASCB, RCEA

incomplete version:

DO NOT QUOTE WITHOUT PERMISSION:

February 15, 2019

Abstract

This paper studies peer effects in heterogeneous networks. In a composite model allowing for peer effects arising from both, local-aggregate and local-average information, we allow for further network heterogeneity, such that peer effects can differ across networks depending on exogenous factors. A novel Instrumental Variables - Minimum Distance estimator (IV-MDE) is proposed, which nests standard homogeneous peer effects model approaches.

Our empirical application is based on network data of 90 school classes of secondary schools in Germany. Our findings indicate that ignoring network heterogeneity leads to misleading interpretations. We show that important effects discussed in the literature on school performance, such as class size and gender composition, operate indirectly through class specific peer effects.

Keywords: Social network, Peer effects, Instrumental Variables, Minimum Distance Estimation, School Achievement

JEL classification: D85, L14, C3, I21,

* Financial support by the Graduate School of Decision Sciences (GSDS), the German Academic Exchange Service (DAAD) and the DFG Centre of Excellence 2117 "Centre for the Advanced Study of Collective Behaviour" (CASCB) (ID: 422037984) is gratefully acknowledged. For helpful comment we like to thank Michael König, Guido Schwerdt, Tobias Kärner and Stephan Schumann. The usual disclaimer applies.

[a] Department of Economics, Universitätsstraße 1, D-78462 Konstanz, Germany. Phone: +49-7531-88-4214, email: Livia.Shkoza@uni-konstanz.de.

[b] Department of Economics, Ludwigstr. 33, D-78462 Munich, Germany. Phone: +49 89-2180-2224, email: Derya.Uysal@econ.lmu.de.

[c] Department of Economics, Universitätsstraße 1, D-78462 Konstanz, Germany. Phone: +49-7531-88-2660, email: Winfried.Pohlmeier@uni-konstanz.de.

1 Introduction

It is an uncontested hypothesis in the social sciences that social interactions shape an individual's behavior and the behavior of groups of individuals in total, whereby social networks have become the utmost popular modeling framework to describe these interactions. In particular, the study of peer effects has become popular, because they summarize compactly and in an easy to interpret fashion the dependence of an individual's behavior on the behavior of others in a socially interactive environment.

The way in which peer behavior actually arises and how it affects an individual's behavior in a given environment is, however, not yet fully understood. A popular assumption is that peer effects reflect norm behavior in the sense that the individuals align their own behavior with the standards of their peer group represented by the mean behavior of the peer group, as deviating behavior from one's own peer group may inflict a loss in terms of the individual's utility. This idea is reflected in the local-average model specification, which uses the mean behavior of the peer group as a predictor for the individual's behavior. Alternatively, one can argue that an individual profits from the pure size of its network and the quality of the peers therein. For instance, if the transmission mechanism between individual and its network is simply the quality of the information the local-aggregate model specification seems to be a reasonable behavioral hypothesis. Recently, Liu, Patacchini, and Zenou (2014) have shown that both specifications can be incorporated in a single composite model as a result of utility maximizing behavior of N agents in a Nash equilibrium.

The econometric approaches adopted to estimate peer effects are strongly driven by the quality and the informational content of the data available. The key issue is to identify endogenous peer effects estimates in the presence of the reflection problem (see for instance Manski (2000) and Angrist (2014)). One strand of the literature uses quasi-experimental settings of the data, which circumvents the endogeneity issue by the quasi-experimental design.

In this paper, we follow the alternative approach of identifying peer effects. It follows the model by Blume, Brock, Durlauf, and Jayaraman (2015) on linear social interactions and leads to an identifiable, structurally interpretable linear relationship between an individual's outcome variable and the peer effect. By exploiting the network structure along the lines of the econometric literature on spatial lags, exogenous variation in the covariates of the second order peers ('peers of the peers') serve as valid instruments to identify endogenous peer effects if the network is sufficiently sparse.

This paper takes a closer look at peer effects in heterogeneous networks. Starting from the composite model, which incorporates the local-aggregate and the local-average peer effect hypothesis, we introduce network specific heterogeneity. This allows us to study, how exogenous network specific factors determine overall peer behavior. Since the model can be derived from utility maximizing behavior, the network specific peer effects have a structural interpretation and provide a different perspective on how heterogeneity affects peer behavior and lastly the outcome variable under investigation. We propose an easy to implement IV - Minimum Distance estimator (IV-MDE), which parametrically nests conventional network based peer effects model.

The empirical evidence on the identification of peer effects exploiting network information is rather scarce due to the limited availability of appropriate network data. It seems fair to state, that so far empirical evidence on the existence and economic relevance of peer effects in networks is rather mixed and refers to various outcome variables and estimation approaches. Using Generalized 2SLS strategy proposed by Kelejian and Prucha (1998), Bramoullé, Djebbari, and Fortin (2009) estimate peer effects on recreational activities for the local-average model using the In-school Add Health data. Using the same data set Liu et al. (2014) study peer effects on effort and sport activities within the composite model framework. They find for study effort, small positive but insignificant local-average and local-aggregate IV-estimates of the peer effects. However, for the GMM estimates they find significant local-average peer effects. With regard to the sports activities, the estimated local-aggregate effect also turns out to be small, but is statistically significant. Boucher, Bramoullé, Djebbari, and Fortin (2014) use the data from the Quebec Government MERS, where the dependent variables are individual scores in standardized tests in Math, Science, French and History. They estimate the local-average model, assuming that a student's reference group consist of class mates only. Based on conditional ML and pseudo conditional ML, they find significant peer effects only for Math, while the IV estimates for the same model are insignificant. None of the studies quoted above take into account network specific heterogeneous peer effects.

In our empirical application, we study the heterogeneity of peer effects on educational achievement using unique network data from 90 school classes of secondary schools in Germany. Our study reveals that ignoring network heterogeneity generally leads to insignificant estimates of peer effects, while accounting for network heterogeneity yields novel insights on how the peer effects operate. In particular, our study provides a better understanding of how the gender composition affects an individual's educational achievement by enhancing peer behavior. Our study also contributes to the long lasting debate on the effect of class size on educational achievement. While many empirical studies concentrate on direct effects of class size on educational achievement within reduced form settings, our more structurally motivated approach, identifies another pathway on how class size affects indirectly educational achievement via peer behavior.

The outline of the paper is as follows. In Section 2 we introduce the composite network model and work out its identification condition. The new minimum distance approach for the estimation of heterogeneous peer effects is introduced in Section 3. In Section 4 we give a description of our network data and discuss further implementation issues. Section 5 contains the major empirical findings, while Section 6 concludes and gives an outlook on future research.

2 The Network Model

Following Liu et al. (2014), in this set up we assume a finite set of N agents, partitioned into L independent networks, where n_l denotes the number of agents in the lth network $(l = 1, ..., L)$. Social connections for network l are recorded in the adjacency matrix $A_l = [a_{ij,l}]$, where $a_{ij,l} = 1$ if agent i is connected with agent j, and $a_{ij,l} = 0$ otherwise. The diagonal elements $a_{ii,l}$ are set to zero. The reference group of some agent i in network l is the set of his peers, and the size of the reference group is the (out)degree $a_{i,l} = \sum_{j=1}^{n_l} a_{ij,l}$.

Define $G_l = [g_{ij,l}]$ the row-normalized adjacency matrix of network l, with elements $g_{ij,l} =$

$a_{ij,l}/a_{i,l}$. By construction we have that $0 \leq g_{ij,l} \leq 1$ and $\sum_{j=1}^{n_l} g_{ij,l} = 1$. Each agent i exerts time or effort $y_{i,l}$ in some activity and $Y_l = (y_{1,l}, ..., y_{n_l,l})'$ is the vector of effort for network l. The agents choose their level of effort to maximize the utility function

$$u_{i,l} = u_{i,l}(y_{i,l}; Y_l, A_l) = \underbrace{\left(\pi_{i,l}^* + \lambda_{1l} \sum_{j=1}^{n_l} a_{ij,l} y_{j,l}\right) y_{i,l}}_{benefit} - \underbrace{\frac{1}{2}\left[y_{i,l}^2 + \lambda_{2l}\left(y_{i,l} - \sum_{j=1}^{n_l} g_{ij,l} y_{j,l}\right)^2\right]}_{cost}, \quad (1)$$

where $\lambda_{1l} \geq 0$ and $\lambda_{2l} \geq 0$.

The benefit from exerting effort level $y_{i,l}$ is a linear function of own effort, with return equal to $\pi_{i,l}^* + \lambda_{1l} \sum_{j=1}^{n_l} a_{ij,l} y_{j,l}$, where $\pi_{i,l}^*$ captures ex ante individual heterogeneity in the return to effort. The aggregate effort of i's peers is $\sum_{j=1}^{n_l} a_{ij,l} y_{j,l}$ with the social multiplier coefficient λ_{1l}. The aggregate effort is heterogeneous in i because individuals have different locations in the network.

The first term of the cost component $y_{i,l}^2$ represents the cost of own effort. Considering the average effort of the peers, $\sum_{j=1}^{n_l} g_{ij,l} y_{j,l}$, as a social norm, the second term $\left(y_{i,l} - \sum_{j=1}^{n_l} g_{ij,l} y_{j,l}\right)^2$ captures the costs of deviating the social norm of the reference group with the social-conformity coefficient λ_{2l}. Note that the cost/benefit parameters λ_{1l} and λ_{2l} are network specific, i.e. the effects of the peers depend on the specific network environment the agent is a part of. In this setting, the total effort of the reference group has a positive impact on the agent's utility, while the distance from the average effort of the peers negatively affects individual utility.

The best reply function from the first-order condition of the utility maximization of individual i is given by:

$$y_{i,l} = \beta_{1l} \sum_{j=1}^{n_l} a_{ij,l} y_{j,l} + \beta_{2l} \sum_{j=1}^{n_l} g_{ij,l} y_{j,l} + \pi_{i,l}, \quad (2)$$

with $\beta_{1l} = \lambda_{1l}/(1 + \lambda_{2l})$ and $\beta_{2l} = \lambda_{2l}/(1 + \lambda_{2l})$, and $\pi_{i,l} = \pi_{i,l}^*/(1 + \lambda_{2l})$. The coefficient β_{1l} captures the local-aggregate endogenous peer effect and it reflects strategic complementarity in efforts. The coefficient β_{2l} captures local-average peer effects and it represents the tendency for conformity. One important detail to note is that $\beta_{1l}/\beta_{2l} = \lambda_{1l}/\lambda_{2l}$, meaning that the relative magnitude of β_{1l} and β_{2l} is the same as that of the social-multiplier coefficient λ_{1l} and the social-conformity coefficient λ_{2l}.

Denote by a_l^{max} the highest degree in the network l. Let $\Pi_l = (\pi_{1,l}, ..., \pi_{n_l,l})'$. Liu et al. (2014) show that if $\beta_{1l} \geq 0$, $\beta_{2l} \geq 0$ and $a_l^{max}\beta_{1l} + \beta_{2l} < 1$, then the network game with payoffs (1) has a unique interior Nash equilibrium in pure strategies given by

$$Y_l = (I_{n_l} - \beta_{1l}A_l - \beta_{2l}G_l)^{-1}\,\Pi_l. \tag{3}$$

In this setting, the higher the effort of the peers, the higher is the individual's utility from exerting own effort (strategic complementarities). Therefore, there is a problem of existence of equilibrium due to the absence of bound on the individual's effort. The condition $a_l^{max}\beta_1 + \beta_2 < 1$ limits the degree of complementarity, and as a result ensuring the existence of an equilibrium. The problem of multiple equilibria is solved by the presence of linear best-reply function and strategic complementarities.

Local-average and local-aggregate specifications

When there is no strategic complementarity effect ($\lambda_{1l} = 0$), the utility function (1) reduces to

$$u_{i,l} = u_{i,l}(y_{i,l}; Y_l, A_l) = \underbrace{\pi^*_{i,l}y_{i,l}}_{benefit} - \underbrace{\frac{1}{2}\left[y_{i,l}^2 + \lambda_{2l}\left(y_{i,l} - \sum_{j=1}^{n_l} g_{ij,l}y_{j,l}\right)^2\right]}_{cost}.$$

The best-reply function of individual i is given by

$$y_{i,l} = \beta_{2l}\sum_{j=1}^{n_l} g_{ij,l}y_{j,l} + \pi_{i,l}, \tag{4}$$

where $\beta_{2l} = \lambda_{2l}/(1 + \lambda_{2l})$ and $\pi_{i,l} = \pi^*_{i,l}/(1 + \lambda_{2l})$. Because the individual effort only depends on average effort of the reference group, this model is referred to as the local-average network model. If $0 \leq \beta_{2l} \leq 1$ the unique interior Nash equilibrium is given by

$$Y_l = (I_{n_l} - \beta_{2l}G_l)^{-1}\,\Pi_l. \tag{5}$$

When there is no social conformity effect ($\lambda_{2l} = 0$), the utility function becomes

$$u_{i,l} = u_{i,l}(y_{i,l}; Y_l, A_l) = \underbrace{\left(\pi^*_{i,l} + \lambda_{1l} \sum_{j=1}^{n_l} a_{ij,r} y_{j,l}\right) y_{i,l}}_{benefit} - \underbrace{\frac{1}{2} y^2_{i,l}}_{cost}.$$

Then, the best-reply function is given by

$$y_{i,l} = \beta_{1l} \sum_{j=1}^{n_l} a_{ij,l} y_{j,l} + \pi_{i,l}, \tag{6}$$

where $\beta_{1l} = \lambda_{1l}/(1 + \lambda_{2l}) = \lambda_{1l}$ and $\pi_{i,l} = \pi^*_{i,l}/(1 + \lambda_{2l}) = \pi^*_{i,l}$. This model is referred to as the local-aggregate network model because the individual effort depends on the aggregate effort of the peer group. If $0 \leq a_l^{max} \beta_l \leq 1$, then the unique interior Nash equilibrium is given by

$$Y_l = (I_{n_l} - \beta_{1l} A_l)^{-1} \Pi_l. \tag{7}$$

The difference between the two models is that in the local-average model individual utility is higher when individual effort is closer to the average effort of the peer group, while in the local-aggregate model individual utility is higher when the agent has more active peers. In the local-average model positions of the agents in the network do not play a role, while in the other one equilibrium effort levels would vary with positions, even if agents are ex ante identical.

Identification

The econometric model has its micro-foundations in the best-reply function of the network game. Define individual heterogeneity

$$\pi_{i,l} = x'_{i,l} \delta_l + \sum_{j=1}^{n_l} g_{ij,l} x'_{j,l} \gamma_l + \eta_l + \epsilon_{i,l},$$

where $x_{i,l}$ is a p-dimensional vector of exogenous variables of agent i in network l, $\epsilon_{i,l}$ is an i.i.d innovation with zero mean and finite variance σ^2, and δ_l, γ_l, η_l are corresponding parameters.

The best reply function (2) leads to the general econometric model:

$$y_{i,l} = \beta_{1l} \sum_{j=1}^{n_l} a_{ij,l} y_{j,l} + \beta_{2l} \sum_{j=1}^{n_l} g_{ij,l} y_{j,l} + \sum_{j=1}^{n_l} g_{ij,l} x'_{j,l} \gamma_l + x'_{i,l} \delta_l + \eta_l + \epsilon_{i,l}, \tag{8}$$

for $i = 1, .., n_l$ and $l = 1, ..., L$. Let $Y_l = (y_{1,l}, ..., y_{n_l,l})'$, $X_l = (x_{1,l}, ..., x_{n_l,l})'$ and $\epsilon_l = (\epsilon_{1,l}, ..., \epsilon_{n_l,l})'$ and ι_{n_l} a $n_l \times 1$ vector of ones. Then 8 gives in matrix form:

$$Y_l = \beta_{1l} A_l Y_l + \beta_{2l} G_l Y_l + +G_l X_l \gamma_l + X_l \delta_l + \eta_l \iota_{n_l} + \epsilon_l. \tag{9}$$

The model coefficients β_{1l} and β_{2l} capture the endogenous effects, where the individual's choice depends on the choices of the peers, γ_l captures contextual effects, where the individual's choice depends on the characteristics of the peer group, and network-specific parameter η_l captures the network fixed effects, where agents in a network might behave similarly because they face similar environmental characteristics or due to similar unobserved individual characteristics (see Manski (1993)). Liu et al. (2014) explain the inclusion of network fixed effects as a partial solution to the self-selection problem. Considering the network formation model as a two-step procedure, in the first step individuals self-select themselves into networks, and then in a second step, they form links within networks based on observable individual characteristics.

If homogeneity of the peer effects is assumed, i.e $\beta_{1l} = \beta_1$, $\beta_{2l} = \beta_2$, $\gamma_l = \gamma$, $\delta_l = \delta$ for $l = 1, ..., L$, the parameters can be estimated using a pooled model defined as follows. Let $diag\{M_j\}$ be the generalized block diagonal matrix in which the diagonal blocks are $k_j \times n_j$ matrices M_j. For a dataset with L groups, define $Y = (Y'_1, ..., Y'_L)'$, $X = (X'_1, ..., X'_L)'$, $\eta = (\eta_1, ..., \eta_L)'$, $\epsilon = (\epsilon'_1, ..., \epsilon'_L)'$, $A = diag\{A_l\}_{l=1}^L$, $G = diag\{G_l\}_{l=1}^L$ and $D = diag\{\iota_{n_l}\}_{l=1}^L$.

$$Y = \beta_1 AY + \beta_2 GY + GX\gamma + X\delta + D\eta + \epsilon \tag{10}$$

To eliminate the fixed effects parameter η, Bramoullé et al. (2009) suggest two types of first differencing. Local differencing means averaging equation (8) over all individual i's peers, and subtract it from i's equation. They call this approach local because it assumes the fixed effect is the same only for i's peers, and not for the whole network. Global differencing means averaging equation (8) over all individuals in i's network and then subtract it from i's equation. In both cases equation (9) is multiplied by a matrix $J = diag\{J_l\}_{l=1}^{L}$. For local differencing $J_l = I_{n_l} - G_l$ and for global differencing $J_l = I_{n_l} - \frac{1}{n_l}\iota_{n_l}\iota_{n_l}'$. Liu et al. (2014) recommend to use global differencing. Because $JD = 0$, the transformed model is given by:

$$JY = \beta_1 JAY + \beta_2 JGY + JX\delta + JGX\gamma + J\epsilon. \tag{11}$$

Identification of the local-average model

The general network model reduces to the local-average model when $\beta_{1l} = 0$:

$$Y_l = \beta_{2l} G_l Y_l + G_l X_l \gamma_l + X_l \delta_l + \eta_l \iota_{n_l} + \epsilon_l, \tag{12}$$

with the reduced form

$$Y_l = (I_{n_l} - \beta_{2l} G_l)^{-1} (G_l X_l \gamma_l + X_l \delta_l + \eta_l \iota_{n_l} + \epsilon_l). \tag{13}$$

Because $J_l(I_{n_l} - \beta_{2l}G_l)^{-1}\iota_{n_l} = 0$ and $(I_{n_l} - \beta_{2l}G_l)^{-1} = \sum_{j=0}^{\infty}(\beta_{2l}G_l)^j$, it follows that

$$\begin{aligned} E[J_l G_l Y_l] &= J_l G_l X_l \delta_l + J_l G_l^2 \big(I_{n_l} - \beta_{2l} G_l\big)^{-1} X_l \big(\beta_{2l}\delta_l + \gamma_l\big) \\ &= J_l G_l X_l \delta_l + \big(J_l G_l^2 X_l + \beta_{2l} J_l G_l^3 X_l + \dots\big)\big(\beta_{2l}\delta_l + \gamma_l\big). \end{aligned}$$

For the identification to work, it has to be that $\beta_{2l}\delta_l + \gamma_l \neq 0$. If $\beta_{2l}\delta_l + \gamma_l = 0$, the $E[J_l G_l Y_l] = J_l G_l X_l \delta_l$ and $E[J_l W_l] = [E[J_l G_l Y_l], E[J_l X_l], E[J_l G_l X_l]]$ does not have full column rank. In this case, the only informative IV for the endogenous effects is $J_l G_l X_l$,which is also a regressor in the structural equation. Here it becomes clear that the perfect collinearity between the mean of the endogenous effect and the contextual effects makes the model unidentified. The exclusion restriction for the identification is provided by the intransitivity of the network, meaning that the peers of the agents' peers are not necessarily the agents' peers. Otherwise, $G_l^2 X$ will be perfectly correlated with X_l and $G_l X_l$. Bramoullé et al. (2009) show that this model is identified given that $I_{n_l}, G_l, G_l^2, G_l^3$ are linearly independent. A possible instrument matrix in this model is given by $Z_l = J_l[X_l, G_l X_l, G_l G_l X_l]$, where $G_l G_l X_l$ represents the mean characteristics of the peers of the peers.

Identification of the local-aggregate model

If $\beta_{2l} = 0$, the network model reduces to local-aggregate model

$$Y_l = (I_{n_l} - \beta_1 A_l)^{-1}(X_l\delta + G_l X_l \gamma + \eta_l \iota_{n_l} + \epsilon_l),$$

implying that the mean of the endogenous variable is

$$E[J_l A_l Y_l] = J_l A_l (I_{n_l} - \beta_1 A_l)^{-1}(X_l\delta + G_l X_l\gamma) + \eta_l J_l A_l (I_{n_l} - \beta_1 A_l)^{-1}\iota_{n_l} \tag{14}$$

The term $A_l(I_{n_l} - \beta_1 A_l)^{-1}\iota_{n_l}$ represents the Katz-Bonacich measure of centrality (Bonacich (1987),Katz (1953)). If the row sums of A_l are not constant, then $J_l A_l (I_{n_l} - \beta_1 A_l)^{-1}\iota_{n_l} \neq 0$, resulting in an additional instrument for the endogenous variable. A possible instruments matrix for the local-aggregate model is $Z_l = J_l[X_l, G_l X_l, A_l X_l]$. Liu et al. (2014) suggest that the identification conditions for the local-aggregate model are weaker than those for the local-average model given in Bramoullé et al. (2009).

Identification of the general model

For the lth network the reduced form is given by

$$Y_l = (I_{n_l} - \beta_1 A_l - \beta_2 G_l)^{-1}(X_l\delta + G_l X_l\gamma + \eta_l \iota_{n_l} + \epsilon_l),$$

implying

$$E[J_l A_l Y_l] = J_l A_l (I_{n_l} - \beta_1 A_l - \beta_2 G_l)^{-1}(X_l\delta + G_l X_l\gamma) + J_l A_l (I_{n_l} - \beta_1 A_l - \beta_2 G_l)^{-1}\iota_{n_l}\eta_l, \tag{15}$$

and

$$E[J_l G_l Y_l] = J_l G_l (I_{n_l} - \beta_1 A_l - \beta_2 G_l)^{-1}(X_l\delta + G_l X_l\gamma) + J_l G_l (I_{n_l} - \beta_1 A_l - \beta_2 G_l)^{-1}\iota_{n_l}\eta_l. \tag{16}$$

The identification condition in this case also relies on some variation in the degrees of some network l and $\eta_l \neq 0$. In this case $J_l A_l (I_{n_l} - \beta_1 A_l - \beta_2 G_l)^{-1}\iota_{n_l}$ can be used as an instrument for $J_l A_l Y_l$ and $J_l G_l (I_{n_l} - \beta_1 A_l - \beta_2 G_l)^{-1}\iota_{n_l}$ for $J_l G_l Y_l$. In their application Liu et al. (2014)

use as instruments $Z_{l1} = J_l[X_l, G_lX_l, A_lX_l, G_lG_lX_l]$ and for the bias-corrected 2SLS estimator they use $Z_{2l} = J_l[X_l, G_lX_l, A_lX_l, G_lG_lX_l, A_l\iota_{n_l}]$. $A_l\iota_l$ is a vector containing the outdegrees of the individuals. The inclusion of this instrument improves asymptotic efficiency and facilitates identification when the standard set of instruments are weak.

3 Minimum Distance Estimation

In the following we introduce our minimum distance estimator for heterogeneous peer effects for the composite model accounting for the endogeneity of the social multipliers. In the first estimation stage the parameters of (9) are estimated by instrumental variables using globally differenced regressors:

$$\tilde{Y}_l = W_l\pi_l + \tilde{\varepsilon}_l, \quad (l = 1, \dots, L),$$

where $\tilde{Y}_l = J_lY_l$ is globally differenced vector of dependent variables and $\pi_l = \left(\beta_{1l}, \beta_{2l}, \delta_l', \gamma_l'\right)'$ is the first stage parameter vector of dimension $2(1 + k_x) \times 1$. The $n_l \times 2(1 + k_x)$ regressor matrix W_l takes the form $W_l = J_l\left[A_lY_l \quad G_lY_l \quad X_l \quad G_lX_l\right]$. As the networks are assumed to be independent, no efficiency gains of a systems regression approach arise over the a simple single equation estimation approach. As the standard set of instruments we use the exogenous variables, their counterparts for the peers for, as well as the covariates of the second order peers as over-identifying instruments, $Z_l = J_l\left[X_l \quad G_lX_l \quad A_lX_l \quad G_l^2X_l\right]$. In a second specification we use the vector of outdegrees as instruments, as suggested in Liu et al. (2014), to account for the weak instruments bias.

In what follows we assume that the network specific peer effects parameters β_{1l} and β_{2l} are endogenous and can be explained by a set of network-specific observable factors:[1]

$$\beta_{jl} = m_l'\beta_j\,, \quad (j = 1, 2),$$

where m_l is a $k_m \times 1$-vector of network specific characteristics and k_x the number of exogenous characteristics contained in X_l and its neighborhood counterpart G_lX_l, respectively. For the

[1] This assumption can be generalized by assuming that β_{jl} can be replaced by a linear predictor representation along the lines of Chamberlain (1984) for panel data models with correlated random effects.

remaining parameters, we assume that they are the same across networks, i.e. $\gamma_l = \gamma$ and $\delta_l = \delta$. The restriction between the first stage reduced form parameter vector π_l and the structural form parameter vector $\theta = \left(\beta_1, \beta_2, \gamma', \delta'\right)'$ is given by:

$$\pi_l = M_l \theta \tag{17}$$

with

$$M_l = \begin{bmatrix} m_l' & 0 & 0 & 0 \\ 0 & m_l' & 0 & 0 \\ 0 & 0 & I_{k_x} & 0 \\ 0 & 0 & 0 & I_{k_x} \end{bmatrix}_{2(1+k_x)\times 2(k_m+k_x)} .$$

Stacking the restrictions between the reduced form parameter π_l and the structural form parameter θ given by (17) into a hyper-system yields:

$$\pi = M\theta \tag{18}$$

where

$$\pi = \begin{bmatrix} \pi_1 \\ \pi_2 \\ \vdots \\ \pi_L \end{bmatrix} \qquad M = \begin{bmatrix} M_1 \\ M_2 \\ \vdots \\ M_L \end{bmatrix}$$

The optimal minimum distance estimator minimizes the weighted quadratic distance between the estimated reduced form parameter vector π from the first stage regression and the structural form parameter:

$$\hat{\theta}_{MD} = \arg\min_{\theta} \left[\hat{\pi} - M\theta\right]' \hat{\Omega}^{-1} \left[\hat{\pi} - M\theta\right] \tag{19}$$

Since the restriction between π and θ is linear the MD estimator can be simply obtained by a generalized least squares regression of $\hat{\pi}$ on M:

$$\hat{\theta}_{MD} = \left(M'\hat{\Omega}^{-1}M\right)^{-1}M'\hat{\Omega}^{-1}\hat{\pi}$$

with $\hat{V}[\hat{\theta}_{MD}] = \left(M'\hat{\Omega}^{-1}M\right)^{-1}$. Note that the optimal weighting matrix is typically of high dimension. However, because of the independence assumption, the optimal weighting matrix is block-diagonal $\hat{\Omega}^{-1} = diag[\hat{\Omega}_l^{-1}]$, where $\hat{\Omega}_l = \hat{V}[\hat{\pi}_l]$. Therefore, the minimum distance estimator can be simply computed as $\hat{\theta}_{MD} = \left(\sum_{l=1}^{L} M_l'\hat{\Omega}_l^{-1}M_l\right)^{-1}\left(\sum_{l=1}^{L} M_l'\hat{\Omega}_l^{-1}\hat{\pi}_l\right)$ even for a large number of networks and large number of covariates.

4 Data

Our empirical study is based on data of the first wave of the *Gymnasiasten-Studie*, a longitudinal survey of about 3000 10th grade students attending an upper secondary school ("Gymnasium") in the German federal state North Rhine-Westphalia (NRW) in the years 1969 and 1970. The students were sampled from 121 classes at 68 upper secondary schools. At the time of the first wave, the students' age was between 13 and 18 with an average of 15 years.

Central to our study is the network information collected in the Sociometric Test of the *Gymnasiasten Studie*. In order to construct the adjacency matrices G_l and A_l for each class we use information on every student's assessment of whom he or she likes in the class based on the question:[2]

> *"In every class there are fellow students, who one likes more in the class than others. Some others one finds pretty unpleasant and that is quite normal. Would you first list the students, who you personally like a lot."*

It is important to note, that attrition is negligible, because every member of every class was surveyed if they were absent on the day of the interview or not. Thus the participation of nearly all class members makes it possible to observe almost the entire social network at the class level. The raw data of the first wave contain information on 3229 students in 121 classes with an

[2] The original question in German is: *"In jeder Klasse gibt es Mitschüler, die man sympathisch findet und die man mehr als andere in der Klasse gut leiden kann. Einige findet man sicher recht unsympathisch, und das ist auch ganz normal so. Würden sie num zunächst einmal die Schüler nennen, die Sie persönlch gut leiden können."*

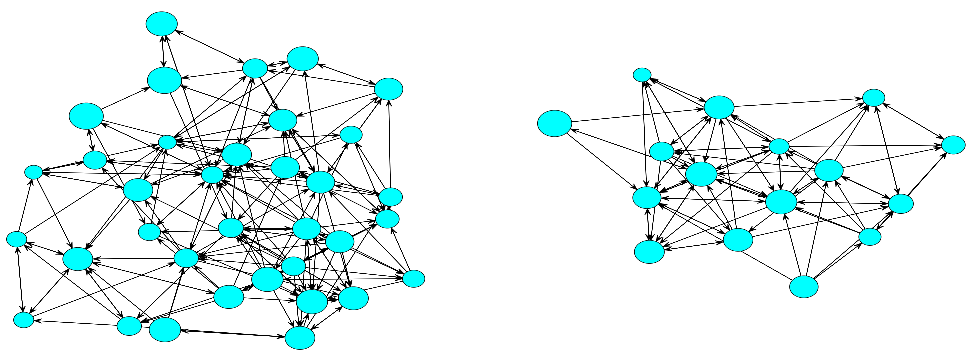
average class size of 26.69 students, with a minimum of 14 and maximum of 39 students. In our empirical study we exclude classes with fewer than 16 student for two reasons. First, the social process in small network are likely to work differently. Secondly, the first-stage estimates of a small networks suffer from low degrees of freedom. Excluding small classes lead to 90 in total with an average class size of 24.93. Table 1 contains summary statistics of the networks.

Table 1: Summary Statistics of the Networks

Variable	Min	Mean	Max	Std. Dev	No nomination
Outdegree, like	0	5.48	29	3.39	15
Indegree, like	0	5.48	22	3.55	56
N = 2244					

The average number of friends a student nominates (outdegree) is 5.48, which indicates that the students take the selection of friends seriously. 15 students do not nominate any friends and 56 students are not nominated as a friend by any other classmate. The distribution of the outdegrees and the indegrees indicates that the networks are sparse such that identification of the network specific peer effects is guaranteed. Figure 1 depicts for illustrative purposes the largest and the smallest network in our sample. Simple eye-balling reveals that the centrality of a student does not necessarily depend on his/her performance. While in the largest network on the left a student with excellent GPA scores receives many nominations, two students with below average performance in the smallest class are central to the network in terms of the number of nominations. But for both classes, there exist peripheral high and low performers.

Figure 1: The largest and the smallest classroom network in the sample



Left: Largest classroom network in the sample with 35 members. Right: smallest classroom network with 16 members. The size of the nodes indicate the student's GPA score, German grading system, smaller nodes indicate better performing students. Source: NRW Gynasiasten-Studie

In our heterogeneous peer effects model local-average and local-aggregate peer effects are allowed to depend on class specific variables. Obvious candidates are class size and the fraction of girls in the class because they are likely to be major determinants of collaborative behavior in the classroom.

As outcome variable, we use the students overall GPA constructed as the average grades over all compulsory and elective courses. In fact, at the time of the survey the choice of different courses within a class was very limited, i.e. all students of the same class basically had to take the same courses. Selection to certain specializations (e.g. languages, mathematics and sciences, humanistic secondary school) took place with the choice of the specific secondary school. The GPA is measured in terms of the German Grading systems with 1 (”very good”) being the best grade 6 (”insufficient”) as the worst grade. Besides the overall GPA we take a closer look at the score in Mathematics and German in order to detect potential differences in peer behavior across subjects.

In order to control for individual heterogeneity we use the the student’s IQ score measured by the Amthauer test, the overall GPA score from the previous school year and the students age. Table A.4 in the Appendix contains a short description of the variables used and the corresponding summary statistics.

5 Empirical Results

In this section, we compare the IV estimates for the composite model with homogeneous peer effects with the ones for the composite model with heterogeneous peer effects based on the new minimum distance IV approach. Although most of the empirical studies focus on peer effects as a result of norm behavior, and therefore favor the local-average model, ex-ante both hypothesis in which way peers affect individual educational achievement are reasonable. For instance, Hoxby (2002) argues, that ’misbehaving students may disrupt the classroom, thereby snapping the teachers’ time and energy.’ In a similar vein knowledge spillovers may work. For both pathways, the local-aggregate model accounting for the number and the quality of the peers seems to be more appropriate.

Because peer effects may also work differently for different school subjects, we study the impact of peer behavior on the overall GPA score as well as on the grades in Math and German as separate measures for the student's educational achievement. As predetermined or exogenous explanatory variable, we use the GPA of the previous year, IQ and age as well as their counterparts for the student's peer group. Because the number of instruments varies between 9 and 13, we include only classrooms of more than 15 students. Table 2 reports the estimation results for the composite model with homogeneous peer effects model with global differencing as given by Eq. (11). For the homogeneous peer effects model we generally do not find any significant peer

Table 2: IV Estimation of the Composite Model: Homogeneous Peer Effects

	GPA		Math Grade		German Grade	
	coef	t-stat	coef	t-stat	coef	t-stat
$\hat{\beta}_1$	0.001	1.27	0.002	1.45	0.002	2.20
$\hat{\beta}_2$	0.619	1.46	0.164	0.34	0.368	0.89
IQ	-0.002	-2.49	-0.019	-9.06	-0.005	-3.50
prev_gpa	0.672	44.27	0.710	19.47	0.711	23.54
age	0.004	0.45	0.059	2.86	-0.036	-2.06
IQ_p	0.001	0.60	0.009	1.10	0.003	1.00
prev_gpa_p	-0.377	-1.18	-0.152	-0.40	-0.199	-0.54
age_p	-0.046	-2.31	-0.011	-0.21	-0.034	-1.63
	N=2246, L=90		N=2245, L=90		N=2244, L=90	

IV estimates of the composite model, standard set of instruments, t-values are based on robust standard errors, the subindex p indicates parameters on exogenous peer variables

effects at conventional significance levels. The only exception is the positive local-aggregate peer effect for the grade in German. Our findings are consistent with Boucher et al. (2014), who also use network information for estimating peer effects on educational achievement for the local-average model. Our findings are also in line with Liu et al. (2014), who find no significant peer effects in the composite model estimating peer effects in study effort.

As expected, the previous GPA is a good predictor of the next period's results, supported by the positive coefficient of this variable. The impact of age is significant in Math and German, where in Math being older implies a worse grade, while in German older students receive better grades. Regarding the characteristics of the friends, only the average age of the friends has a positive (negative coefficient), significant impact on the final GPA of the individuals.

Table 3 summarizes the estimation results for the composite model with heterogeneous peer effects based on the IV-MDE approach with globally differenced variables. The heterogeneity of the peer effects depends on class size and the share of girls in class. Note, that both variables are centered around the mean, so that the intercept terms reflect the aggregate and the average peer effects for a class of average size and average gender share. Interestingly, for the GPA score both the local-aggregate and local-average peer effects become significant. However, for the two other achievement variables, we find positive and significant local-average peer effects at the 10 % level for the Math grade and at the 5 % level for the German grade. Particularly, the grade in German is strongly influenced by norm behavior as represented by the local-average peer effect. The negative effect of class size on the local-average peer effect for the GPA indicates that in larger classes the peer effect is generally smaller, i.e. the enhancing contribution of peer behavior diminishes with larger classes. Based on the estimates given in Table 3, the local-average peer effect for the GPA score vanishes completely with a class size of 36 students, which is above the maximum class size in our sample. We therefore conclude, that this peer effect is generally present, but maybe rather small for larger classes. It is important to note, that this class size effect is novel in the literature as it operates through peer behavior. This class size effect through peer behavior operates in addition to a potential direct class size on a student's performance, which is traditionally under consideration in studies on educational achievement. Our approach accounts for the direct effect by global differencing, so that the class size effect on peer behavior represents a second channel on the impact of class size on educational achievement.

The class size effects measured in the literature differ from ours, as other studies measure the direct effect of class size on achievement variables. Angrist and Lavy (1999) use the Maimonides rule as an instrument for class size, to estimate the effect of class size on the mean outcome of the class. Their study reports a negative and significant class size effect for Math and reading in the fifth grade, and these effects are smaller than those found in the Tennessee STAR experiment. Hoxby (2000) on the other hand relies on longitudinal variation in the population to identify the effects of class size on student achievement, and finds no significant effect of class size on achievement.

Table 3: IV-MD estimation of the Composite Model, Heterogeneous Peer Effects

	GPA		Math Grade		German Grade	
	coef	t-stat	coef	t-stat	coef	t-stat
Intercept_1	0.002	3.63	0.002	1.52	0.001	1.26
class_ size_1	-0.000	-0.87	0.000	0.87	0.000	1.10
frac_girls_1	-0.001	-0.79	0.000	0.15	-0.002	-1.02
Intercept_2	0.111	2.74	0.086	1.92	0.206	4.92
class_size_2	-0.010	-2.78	0.007	0.98	0.006	1.06
frac_girls_2	0.080	2.02	-0.046	-0.63	-0.071	-1.09
IQ	-0.001	-2.78	-0.020	-13.22	-0.003	-2.94
prev_gpa	0.693	71.07	0.707	25.21	0.748	33.73
age	0.003	0.61	0.060	3.90	-0.031	-2.52
IQ_p	0.001	0.53	0.009	2.88	0.004	1.51
prev_gpa_p	-0.002	-0.07	-0.075	-1.16	-0.076	-1.32
age_p	-0.024	-3.70	-0.013	-0.68	-0.029	-2.13
	N=2246, L=90		N=2245, L=90		N=2244, L=90	

IV-MD estimates of the composite model, standard set of instruments, homoskedastic errors at the class level, subindex p indicates parameters on exogenous peer variables

Via local-average peer effect, the gender composition of classes has a significant positive effect on the GPA score, i.e. the larger the fraction of girls in the class the stronger the local-average peer effect. However, we cannot detect this effect for the two subjects German and Math. Similar to reasoning put forward for the class size effect, the gender composition effect operates through general peer behavior in class. This has to be distinguished from the gender composition effects estimated at the individual level, where the gender composition of an individual's peer group is measured as in studies by Feld and Zölitz (2017) and Brenøe and Zölitz (2018). They estimate reduced form equations with interaction terms between the proportion of female peers and gender dummies ignoring network information.

The coefficients on IQ and and the previous GPA have the expected sign and are significant for all achievement measures. Not very surprising, the GPA of the previous school year is a very good predictor of the current performance. Students with a higher IQ also perform better better. However, the IQ effect is rather small. Even for Math, where the IQ effect is the largest, the grade of a student with the highest IQ compared to a student with mean IQ is only 0.55 grade points better. Accounting in addition for the multiplier effect through peer behavior increases the effect in absolute value to 0.61 grade points.

Similar to the results obtained for the homogeneous peer effects model, the effect of age on educational achievement varies across the different the performance measures. Younger students perform better in Math, but they perform worse in German. For the exogenous peer effects we observe rather different effects depending on the subject under consideration. The older a student's peer group the better are the overall grades and the grades in German, while we do not find any peer-age effect for mathematics. Interestingly, we find a "Big Fish Small Pond" effect for Math. Students with lower IQ-peers do better in Math than those with higher IQ peers.

Robustness Checks

We also estimate the local-average and local-aggregate specification separately, in order to avoid multicollinearity problems (see Table A.5 in the Appendix). However, contrary to Liu et al. (2014) in their study, this strategy does not lead to an increase in precision of the estimates for β_1 or β_2. The local-average peer effect remains insignificant for all achievement measures, while β_1-estimates are similar in sign, size and significance to the ones obtained for the composite model presented in Table 2.

Our IV-MDE estimates are robust to the choice of the variance-covariance estimator. We also estimated the composite, heterogeneous peer effects model by IV-MDE using heteroskedasticity robust variance-covariance estimates in the first estimation stage for each class stage to construct the optimal weighting matrix. The corresponding MD estimates do not differ substantially from the ones presented in Table 3, which are based on class-specific heteroskedasticity only.

6 Conclusions

This paper takes a closer look at peer effects in heterogeneous networks. Starting from the composite network model, which incorporates the local-aggregate and the local-average peer effect hypothesis, we introduce network specific heterogeneity. This allows us to study, in which way exogenous network specific factors determine overall peer behavior. Identification of the peer effects is achieved by using exogenous information on the peers of the peers. Since the model can be derived from utility maximizing behavior, the network specific peer effects have a structural interpretation and provide a different perspective on how heterogeneity affects peer behavior and lastly the outcome variable under investigation. An easy to implement IV - Minimum Distance estimator (IV-MDE) is proposed, which parametrically nests conventional network based peer effects model.

In our empirical application based on a unique network data set of 90 school classes of secondary schools in Germany, we show that accounting for network heterogeneity yields interesting new insights. In our study we find clear evidence that peer behavior in schools is driven by social norms as captured by the local average model. Moreover, we show that ignoring network heterogeneity generally leads to insignificant peer effects, while accounting for network heterogeneity yields to significant estimates of the peer effects.

Studying peer effects heterogeneity in a structural framework leads to interesting novel interpretations in which way class size and gender composition affect educational achievement via peer behavior. In particular, we find that the class size effect on educational achievement obtained from reduced form specifications may simply swallow indirect effects operating through differences in peer behavior. In our study, we find clear evidence that for smaller classes the (positive) peer effect is stronger. For the largest classes in our sample the peer effect of class size is small but still remains positive. Moreover, the gender composition of classes also affects the GPA via peer behavior. The larger the proportion of girls in class, the stronger the peer effect.

Future versions of this paper will contain more robustness checks concerning the choice of the instrument set, the weakness of the instruments and improvements based on finite sample bias corrections of the first stage IV estimates. An interesting path for future research is to study the robustness of our IV-MDE compared to a Maximum Likelihood and a GMM approach for the complete set of networks. Our findings on the relevance of peer behavior using network information are conditional on a given network. The instrumentation using information on the peers of the peers solves the endogeneity problem to obtain consistent peer effects parameter.

Future research should also take into account the formation of peers effects, i.e. the endogenous choice of the network nodes in a dynamic context.

References

Angrist, J. and V. Lavy (1999): "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement," Quarterly Journal of Ecoconomics, 114, 533 – 575.

Angrist, J. D. (2014): "The Perils of Peer Effects," Labour Economics, 30, 98–108.

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

Bonacich, P. (1987): ""Power and centrality: A family of measures"," American Journal of Sociology, 92, 1170–1182.

Boucher, V., Y. Bramoullé, H. Djebbari, and B. Fortin (2014): "Do Peers Affect Student Achievement? Evidence from Canada Using Group Size Variation," Journal of Applied Econometrics, 29, 91–109.

Bramoullé, Y., H. Djebbari, and B. Fortin (2009): "Identification of peer effects through social networks," Journal of Econometrics, 150, 41 – 55.

Brenøe, A. A. and U. Zölitz (2018): "Exposure to More Female Peers Widens the Gender Gap in STEM Participation," Tech. Rep. Working Paper No. 285, University of Zurich, Department of Economics,.

Chamberlain, G. (1984): "Panel Data," in Handbook of Econometrics, Vol. II, ed. by Z. Griliches and M. D. Intriligator, Elsevier Science BV, 1247–1318.

Feld, J. and U. Zölitz (2017): "Understanding Peer Effects: On the Nature, Estimation, and Channels of Peer Effects," Journal of Labor Economics, 35, 387–428.

Hoxby, C. M. (2000): "The Effects of Class Size on Student Achievement: New Evidence from Population Variation," The Quarterly Journal of Economics, 115, 1239–1285.

——— (2002): "The Power of Peers: How Does the Makeup of a Classroom Influence Achievement," Education Next, Summer, 57 –63.

Katz, L. (1953): “”A new status index derived from sociometric analysis”,” Psychometrika, 18, 39–43.

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

Liu, X., E. Patacchini, and Y. Zenou (2014): “”Endogenous peer effects: local aggregate or local average?”,” Journal of Economic Behavior and Organization, 39–59.

Manski, C. F. (1993): “”Identification of endogenous social effects: The reflection problem”,” The Review of Economic Studies, 60, 531–542.

——— (2000): “”Economic analysis of social interactions”,” Journal of Economic Perspectives, 14, 115–136.

A Appendix

Table A.4: Summary Statistics of the Variables

Variable	Min	Mean	Max	Std. Dev
prev_gpa	1.40	3.19	4.67	0.49
GPA	1.54	3.15	4.11	0.45
Math Grade	1	3.44	6	0.92
German Grade	1	3.40	5.00	0.74
IQ	9.00	40.13	68.00	9.16
age	13	15.37	18	0.87
classize	16	24.93	35	4.67
frac_girls	0	0.46	1	0.43
Outdegree	0	5.48	29	3.39
Indegree	0	5.48	22	3.55
N=2244				

Table A.5: IV Estimation of the Local-Aggregate and -Average Model: Homogeneous Peer Effects

	Local-aggregate						Local-average					
	GPA		Math Grade		German Grade		GPA		Math Grade		German Grade	
	coef	t-stat	coef	t-stat	coef	t-stat	coef	t-stat	coef	t-stat	coef	t-stat
$\hat{\beta}$	0.001	1.69	0.002	1.44	0.003	2.30	0.521	1.21	0.156	0.32	0.419	0.97
IQ	-0.002	-2.83	-0.019	-9.26	-0.005	-3.53	-0.002	-2.59	-0.019	-9.08	-0.006	-3.54
prev_gpa	0.677	47.26	0.710	19.54	0.717	24.06	0.673	44.78	0.711	19.50	0.711	23.41
age	-0.002	-0.29	0.060	2.94	-0.041	-2.57	0.003	0.29	0.058	2.80	-0.037	-2.09
IQ_p	0.002	1.05	0.007	1.70	0.004	1.21	0.001	0.64	0.009	1.04	0.003	0.93
prev_gpa_p	0.080	3.02	-0.027	-0.38	0.119	2.08	-0.301	-0.93	-0.134	-0.35	-0.229	-0.59
age_p	-0.019	-2.23	0.006	0.28	-0.024	-1.27	-0.041	-2.06	-0.010	-0.17	-0.035	-1.66
	N=2246, L=90		N=2245, L=90		N=2244, L=90		N=2246, L=90		N=2245, L=90		N=2244, L=90	

IV estimation of the local-aggregate and -average model, t-values are based on robust standard errors, subindex p indicates parameters on exogenous peer variables

Table A.6: IV Estimation of the Composite Model: Homogeneous Peer Effects

	GPA		Math Grade		German Grade	
	coef	t-stat	coef	t-stat	coef	t-stat
$\hat{\beta}_1$	0.001	1.44	0.002	1.44	0.003	2.24
$\hat{\beta}_2$	0.344	1.04	0.248	0.53	0.214	0.60
IQ	-0.002	-2.65	-0.019	-9.06	-0.005	-3.52
prev_gpa	0.674	45.87	0.711	19.40	0.714	23.89
age	0.001	0.15	0.059	2.82	-0.038	-2.23
IQ_p	0.001	0.81	0.010	1.28	0.004	1.09
prev_gpa_p	-0.174	-0.70	-0.216	-0.59	-0.067	-0.21
age_p	-0.034	-2.05	-0.020	-0.38	-0.030	-1.48
	N=2246, L=90		N=2245, L=90		N=2244, L=90	

IV estimation of the composite model, bias correction instruments, t-values are based on robust standard errors,subindex p indicates parameters on exogenous peer variables

Table A.7: MD Estimation of the Composite Model, Heterogeneous Peer Effects

	GPA		Math Grade		German Grade	
	coef	t-stat	coef	t-stat	coef	t-stat
Intercept_1	0.002	3.51	0.002	1.48	0.001	1.32
classize	-0.000	-0.75	0.000	1.02	0.000	1.15
frac_girls	-0.001	-0.71	0.000	0.05	-0.002	-1.00
Intercept_2	0.086	2.19	0.072	1.64	0.175	4.29
classize	-0.011	-3.04	0.007	0.98	0.007	1.26
frac_girls	0.061	1.58	-0.041	-0.57	-0.078	-1.23
IQ	-0.001	-2.88	-0.020	-13.25	-0.004	-3.12
prev_gpa	0.693	72.41	0.706	25.35	0.749	34.11
age	0.002	0.44	0.059	3.88	-0.032	-2.66
IQ_p	0.001	0.57	0.009	2.87	0.004	1.61
prev_gpa_p	0.018	0.51	-0.061	-0.95	-0.049	-0.86
age_p	-0.021	-3.29	-0.013	-0.71	-0.029	-2.09
	N=2246, L=90		N=2245, L=90		N=2244,L=90	

MD estimation of the composite model, bias correction instruments, homoskedastic errors at the class level, subindex p indicates parameters on exogenous peer variables

Table A.8: MD Estimation of the Local-Aggregate and -Average Model: Heterogeneous Peer Effects

	Local-Aggregate						Local-Average					
	GPA		Math Grade		German Grade		GPA		Math Grade		German Grade	
	coef	t-stat	coef	t-stat	coef	t-stat	coef	t-stat	coef	t-stat	coef	t-stat
Intercept	0.002	4.406	0.002	1.463	0.002	1.709	0.168	3.71	0.139	2.75	0.237	5.17
classize	-0.000	-2.112	0.000	0.994	0.000	0.579	-0.010	-2.71	0.005	0.71	0.006	0.98
frac_girls	-0.001	-0.732	-0.001	-0.248	-0.002	-0.857	0.053	1.30	-0.058	-0.71	-0.026	-0.37
IQ	-0.002	-3.123	-0.019	-12.888	-0.004	-3.272	-0.001	-2.10	-0.020	-12.28	-0.004	-2.97
prev_gpa	0.694	70.647	0.707	25.291	0.748	33.492	0.696	66.40	0.684	22.52	0.746	31.76
age	0.001	0.153	0.059	3.863	-0.038	-3.097	0.001	0.19	0.055	3.36	-0.032	-2.50
IQ_p	0.001	1.153	0.008	2.678	0.004	1.579	0.000	0.32	0.010	2.98	0.003	1.23
prev_gpa_p	0.077	3.990	-0.012	-0.219	0.096	2.204	-0.029	-0.72	-0.081	-1.15	-0.088	-1.43
age_p	-0.020	-3.378	-0.002	-0.112	-0.018	-1.324	-0.027	-3.90	-0.025	-1.25	-0.026	-1.78
	N=2246, L=90		N=2245,L=90		N=2244,L=90		N=2246,L=90		N=2245,L=90		N=2244,L=90	

IV-MDE of the local-aggregate and -average model, homoskedastic erros at the class level, subindex p indicates parameters on exogenous peer variables