Carnegie Mellon University
Department of Statistics & Data Science

Advanced Data Analysis Project

Why do Students Avoid Each Other? Investigating Positive and Negative Network Effects

Sara Colando[1]

December 5, 2025

Advisors:
Nynke Niezink[1]
Eva Jaspers[2]

[1] Department of Statistics & Data Science, Carnegie Mellon University
[2] Department of Sociology, Utrecht University

Abstract

Negative ties can greatly influence both positive network dynamics and the formation of beliefs and behaviors in social networks. Examples of negative ties include avoidance, dislike, bullying, and aggression. While there is a broad literature on bullying, less is known about what drives people in a group to avoid each other. Yet, avoidance is a particularly interesting type of relation, since it is often more frequent and less socially complex than aggression or bullying. In this project, we study structural dependencies in classroom avoidance networks and how a student's tendency to avoid others and be avoided relates to their position in the friendship network using data from 193 Dutch classrooms. Fitting an exponential random graph model (ERGM), we find that avoidance networks are highly associated with friendship networks and exhibit strong reciprocity and degree-related endogenous effects. Finally, our conclusions based on the average marginal effects of ERGM parameters do not differ substantively from those based on the original ERGM parameter estimates, contrary to previous findings.

Contents

1 Introduction

While social network studies have traditionally focused on positive ties between individuals (e.g., friendship or cooperation), negative ties can substantially influence both positive network dynamics and the formation of beliefs and behaviors in social networks. Many classical social network theories, such as balance theory (Heider, 1946; Cartwright and Harary, 1956), involve negative ties. Balance theory posits that people strive for cognitive consistency, or *balance*, in their social relationships (Cartwright and Harary, 1956). For example, suppose we had a triad consisting of three actors, a, b, and c, where a is friends with b, and b dislikes c. Then, holding these ties fixed, the triad would be *unbalanced* if a is also friends with c, since a would be friends with someone their friend b dislikes. Conversely, the triad would be *balanced* if a dislikes c, which is consistent with their friend b disliking c. Additionally, negative ties have been found to have a stronger effect on psychological and behavioral outcomes than positive ties, particularly in the workplace (Labianca and Brass, 2006). More recently, accounting for negative ties has helped identify social media users who disproportionately propagate polarizing narratives online (Candellone et al., 2025).

Examples of negative ties include dislike, bullying, aggression, and avoidance. While there is a broad literature on bullying and aggression (Huitsing and Veenstra, 2012; Huitsing et al., 2014; Wójcik and Flak, 2019), less is known about what drives individuals to avoid each other. Yet, avoidance is a particularly interesting type of negative relation because it is often more frequent and less socially complex than bullying and aggression (Kros et al., 2021; Toroslu and Jaspers, 2022). For instance, Wójcik and Flak (2019) find that bullies are often also friends, or *frenemies*, with their victims, making it a very complex social relationship.

To this end, two recent papers have examined avoidance networks in classroom settings. Using longitudinal network data from nine classrooms in two Dutch secondary schools, Kros et al. (2021) compare the evolution of classroom antipathy (dislike), avoidance, and aggression networks. They find that avoidance tends to be reciprocated among classmates and that students tend to avoid classmates of a different gender over those with the same gender. With the same longitudinal network data, Toroslu and Jaspers (2022) study avoidance closure in balanced triad by analyzing the co-evolution of classroom avoidance and friendship networks. They find that students are more likely to avoid friends of classmates they have already avoided. Moreover, in at least one classroom, students avoided the same classmates as their friends (Toroslu and Jaspers, 2022).

Like Kros et al. (2021) and Toroslu and Jaspers (2022), our project focuses on avoidance among classmates. In particular, we use cross-sectional network data collected during Netherlands Wave 1 of the Children of Immigrants Longitudinal Survey in Four European Countries (CILS4EU) to study the structural dependencies in classroom avoidance networks and how a student's tendency to avoid others and be avoided relates to their position in the class's friendship network. Our project extends the work from Kros et al. (2021) and Toroslu and Jaspers (2022) in three ways. First, we account for global network configurations in the classroom friendship networks, such as students' connectedness. Second, we control for lower-level, non-triadic structures between the avoidance and friendship networks, such as the tie overlap between the two networks. Third, we use a larger sample of classrooms to increase the generalizability of our findings.

Fitting an exponential random graph model (ERGM) to 193 Dutch classroom avoidance networks, we find that classroom friendship networks are highly associated with classroom avoidance networks, and that classroom avoidance networks exhibit strong reciprocity and degree-related endogenous effects. We do not find sufficient evidence that gender influences students' avoidance tendencies. In addition, we improve the existing implementation of average marginal effects (AMEs) for ERGMs by aligning the calculation of marginal effects for factor covariates with the standard approach for logistic regression models (Mood, 2009; Perraillon and Hedeker, 2025). Contrary to recent findings by Duxbury and Wertsching (2023), our conclusions based on the AMEs of our block diagonal ERGM parameters do not differ substantively from those based on our original parameter estimates.

2 Data

2.1 Participants and procedures

The data for this project comes from Netherlands Wave 1 of the Children of Immigrants Longitudinal Survey in Four European Countries (CILS4EU) (Kalter et al., 2017). It primarily consists of 14-year-old students and was collected during the 2010-2011 school year. Since a primary goal of the CILS4EU project was to study the structural, social, and cultural integration of children of immigrants with their ethnic-majority peers, schools were randomly selected based on school size and the proportion of students with an immigrant background (Kalter et al., 2017). In particular, schools with a high proportion of immigrants in the target grade were oversampled to ensure that a sufficient number of students with an immigrant background participated in the project (Kalter et al., 2017). In the schools that consented to participate in the study, at least two classrooms with mainly 14-year-old students were randomly selected. Students within these selected classes were then asked to complete surveys on their background, academic achievement, and self-reported relationships with their classmates. Parents and teachers also completed surveys to gain more insights into students' backgrounds and the classroom learning environment (Kalter et al., 2017).

After filtering out classrooms with low survey participation rates (less than 50%) and very small class sizes (under 10 students), our data includes 4,076 students from 193 classrooms across 98 Dutch schools. The 193 classrooms comprise four school tracks: Brugklas, VMBO, Havo, and VWO/ES/IB. The VMBO (pre-vocational) track accounts for 123 of the 193 classrooms (63.7%). Meanwhile, 36 (18.7%) classrooms are from the Havo (senior general) track, and 33 (17.1%) classrooms are from the VWO/ES/IB (pre-university) track. Only one (0.5%) is from a Brugklas classroom, which is a bridge year between primary and secondary education in the Netherlands.

Table 1 summarizes the characteristics across the 193 classrooms in our data. On average, there are 21.1 students per classroom, and 2.0 of them did not participate in CILS4EU. An average of 46.0% of students identify as boys, and 38.8% of students have an immigrant background (i.e., have a country of origin other than the Netherlands). Across all classrooms, 61 students (1.5%) did not report their gender. Given that this is a small proportion of our data, we exclude these individuals from our analysis, leaving us with 4,015 students.

Table 1: Summary statistics of class-level characteristics. The mean is the average across the 193 classrooms, and the standard deviation (SD) measures the variation in the given characteristic across classes.

Class characteristics	**Mean**	**SD**
Total students	21.1	4.8
Did not participate in CILS4EU	2.0	1.8
Percent male	46.0	20.5
Percent immigrants	38.8	27.8

2.2 Measures

We measure *avoidance* in each classroom using the CILS4EU question: "Who would you not want to sit by?" Students could nominate up to five classmates for this question. We code nominations as 1 and non-nominations as 0 to construct a binary, directed avoidance network in each classroom. Throughout this paper, we denote the avoidance network for classroom k as $y^{(k)}$ where:

$$y_{ij}^{(k)} = \begin{cases} 1 & \text{if student } i \text{ avoids classmate } j \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

We also measure *friendship* in each classroom since we are interested in how a student's position in the friendship network influences their avoidance tendencies. We measure friendship with two CILS4EU questions: "Who are your best friends in the class?" and "Who do you often spend time with outside of school?" Students could nominate up to five classmates for the former and as many classmates as they liked for the latter. We consider a student to nominate a classmate as a friend if the student nominated the classmate in either question. We code friend nominations as 1 and non-nominations as 0 to construct a binary, directed friendship network in each classroom. By aggregating students' nominations to the two questions, we better capture all of a student's friends in the class.

2.2.1 Missingness

Not every student completed all the self-reported social relationships questions. 208 (5.1%) students did not complete the questions about which classmates they would like to avoid. Meanwhile, 135 (3.3%) students did not complete the questions about which classmates are their top five best friends, and 134 (3.3%) students did not complete the questions about which classmates they often spend time with outside of school. In our analysis, we treat students who did not complete a given set of questions as having no nominations for that social relationship category. That is, we impute them as having zero outgoing ties in the relevant classroom network. We depict the overlap in tie missingness in the Appendix (Figure A.1).

2.3 Exploratory data analysis

Figure 1 depicts avoidance outdegree and indegree distributions across all 4,015 students. While students could nominate up to five classmates as individuals in the questionnaire, most students do not avoid any of their classmates or only avoid one of them (Figure 1A). Moreover, most students are not avoided by any of their classmates, but a few students are avoided by several of their classmates (Figure 1B).

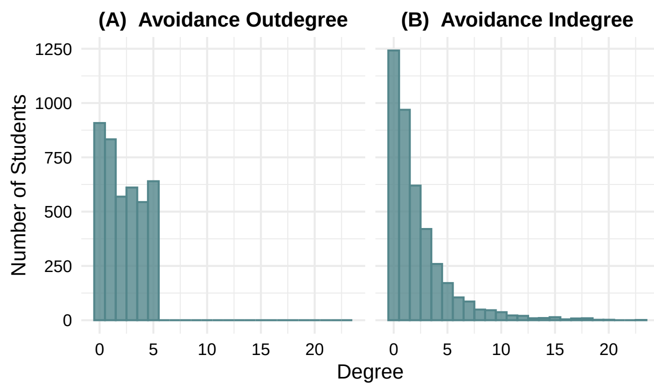


Figure 1: (A) distribution of avoidance outdegree (i.e., the number of classmates the student avoids). (B) distribution of avoidance indegree (i.e., the number of classmates who avoid the student).

Table 2 summarizes key descriptives for the 193 classroom avoidance networks. The avoidance networks have an average density of 0.11 across all the classes in our dataset. Further, the avoidance networks exhibit an average reciprocity of 0.191 and an average transitivity of 0.163. We also find that, on average, 7.4% of students in a classroom are isolates in the sense that they are neither avoided by nor avoid any of their classmates.

As expected, we observe very little overlap between the classroom friendship and avoidance networks: on average, only 2.3% of ties in either network are present in both. We also examine the Jaccard index between the avoidance and transposed friendship networks within a given classroom. The transposed friendship network measures which classmates consider the student a friend. Thus, a tie from classmate j to student i in the transposed friendship network if and only if j nominates i as a friend. We observe that, on average, only 2.7% of ties in the classroom avoidance or transposed friendship networks are present in both networks.

Clustering is also common in classroom friendship networks: students may form close friendship groups and be weakly connected to others in their class (González et al., 2007; Goodreau et al., 2009). In the extreme case, the friendship network could consist of multiple disconnected components. We posit that students are *less* likely to avoid classmates in the same component of the friendship network, and that students in the largest connected component of the friendship network establish the classroom norms (e.g., acceptable social behavior). Hence, being outside the largest connected component of the friendship network means that neither the student nor any of their friends participate in shaping classroom norms. Consequently, we anticipate that less-connected students have different avoidance tendencies than those in the largest connected component of the friendship network. To this end, we identify weakly connected components of each classroom's friendship network. Two students are considered weakly connected in the friendship network if there is an undirected path (of any length) between them. We observe that most students (91.6%) tend to be in the largest weakly connected component of their class's friendship network.

Table 2: Summary statistics for classroom avoidance networks. The mean is the average across the 193 classrooms, and the standard deviation (SD) measures the variation in the given descriptive statistic across the classes.

Descriptive	**Mean**	**SD**
Density	0.109	0.033
Reciprocity	0.191	0.110
Transitivity	0.163	0.113
Isolates	0.074	0.082
Sends at least one tie given:		
Student identifies as a boy	0.810	0.156
Student identifies as a girl	0.792	0.199
Receives at least one tie given:		
Student identifies as a boy	0.735	0.181
Student identifies as a girl	0.683	0.207
Krackhardt's E-I index for gender	0.044	0.045
In largest weakly connected component of class's friendship network	0.916	0.142
Jaccard index:		
Avoidance and friendship	0.023	0.029
Avoidance and transposed friendship	0.027	0.026

Finally, we examine the gender-related descriptives across the avoidance networks in Table 2. We find that boys and girls have approximately the same average probability of avoiding at least one of their classmates (81.0% and 79.2%, respectively). However, girls are less likely to be avoided than boys: on average, 73.5% of boys are avoided by at least one classmate, compared to 68.5% of girls. In line with Kros et al. (2021), we also observe that students tend to avoid classmates with a different gender over those with the same gender (the mean Krackhardt's E-I index for gender is 0.044). However, this tendency is not consistent across all classes. In some classes, students are more likely to avoid their same-gender peers, given that the standard deviation of Krackhardt's E-I Index for gender (0.045) is larger than the mean (0.044).

3 Methods

3.1 Exponential family random graph models (ERGMs)

We model the structural dependencies in classroom avoidance networks using an exponential random graph model (ERGM). ERGMs treat the observed network y as a single realization of a random network. Let $\mathbf{Y}$ denote this random network, where $Y_{ij} = 1$ if actor i sends a tie to actor j, and $Y_{ij} = 0$ otherwise. We assume no self-loops, that is, $Y_{ii} = 0$ for all actors i. In an ERGM, the conditional probability of observing y follows an exponential family distribution (Frank and Strauss, 1986; Wasserman and Pattison, 1996; Anderson et al., 1999):

$$\mathbb{P}_{\theta,\mathcal{Y}}(\mathbf{Y} = y \mid \mathbf{X}) = \frac{\exp(\theta' g(y, \mathbf{X}))}{\kappa(\theta, \mathcal{Y})} \tag{2}$$

Here, $\mathbf{X}$ denotes the covariate matrix, which can consist of both actor-level and dyad-level attributes. The vector $g(y, \mathbf{X})$ contains the sufficient statistics, which include counts for network configurations or a function of these counts within the observed network y, as well as the covariates specified in $\mathbf{X}$. The vector θ is the associated coefficient vector and measures the relative importance of each sufficient statistic to the conditional probability of observing y (Lusher et al., 2012). Finally, $\kappa(\theta, \mathcal{Y})$ is the normalizing constant, which computes the sum of the numerator over all possible networks in the sample space $\mathcal{Y}$ in order to make the outcome a valid probability.

In Equation (2), the outcome is the conditional probability of observing the entire network y. However, ERGMs can also be expressed as a tie-level model, where the outcome is the conditional log-odds of actor i sending a tie to actor j, given the covariate matrix $\mathbf{X}$ and holding all other ties in the network Y_{-ij} fixed:

$$\log\left(\frac{\mathbb{P}(Y_{ij}=1 \mid \mathbf{X}, Y_{-ij})}{\mathbb{P}(Y_{ij}=0 \mid \mathbf{X}, Y_{-ij})}\right) = \theta'\delta^{+}_{ij}(y, \mathbf{X}) \tag{3}$$

The vector $\delta^{+}_{ij}(y, \mathbf{X})$ is the change statistic vector, which measures how much the sufficient statistic vector changes when the focal tie from i to j exists versus when it does not. From Equation (3), we can interpret each coefficient in an ERGM as the change in the conditional log-odds of a tie associated with a one-unit increase in the corresponding statistic (Wasserman and Pattison, 1996).

3.1.1 Block diagonal ERGMs

Since we have 193 classroom avoidance networks, rather than a single one, we cannot directly fit an ERGM to our data. Instead, we use a block diagonal ERGM to accommodate the fact that we have multiple networks. By using a block diagonal ERGM, we treat each classroom avoidance network as an independent block in a larger *supernetwork*. Let y^* denote the supernetwork constructed from the classroom avoidance networks, $y^{(1)}, y^{(2)}, \ldots, y^{(193)}$. The adjacency matrix for y^* can be written as a block diagonal matrix, with the adjacency matrices associated with each classroom avoidance network on the main diagonal and with the off-diagonal elements being structural zeroes since students could only nominate classmates they would like to avoid in the CILS4EU questionnaire:

$$y^* = \begin{bmatrix} y^{(1)} & 0 & \cdots & 0 \\ 0 & y^{(2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & y^{(193)} \end{bmatrix} \tag{4}$$

We can then use an ERGM to model the conditional probability of observing the entire supernetwork y^* or the conditional log-odds of a tie within it. Importantly, since ERGMs are parametrized by a single coefficient vector θ, by fitting a block diagonal ERGM, we assume that the ERGM coefficients are shared across all 193 classroom avoidance networks (Vega Yon et al., 2021). We elaborate on the validity of this assumption in our discussion (Section 5).

3.2 Average marginal effects (AMEs) for ERGMs

In addition to reporting the coefficient estimates, we report the average marginal effects (AMEs) for each block diagonal ERGM parameter, following the recommendations of Duxbury (2021) and Duxbury and Wertsching (2023), who suggest that AMEs improve interpretability and reduce bias in ERGM results. Broadly, AMEs measure the expected change in the predicted outcome for a change in a given covariate, holding all else constant (Perraillon and Hedeker, 2025). Since AMEs are on the outcome scale rather than the link scale, they are often more interpretable than the raw coefficient estimates for nonlinear regression models (Mood, 2009).

Since the outcome of a tie-level ERGM is the conditional log-odds of actor i sending a tie to actor j (see Equation (3)), AME for ERGM parameters are on the probability scale. We can compute the AME for some covariate x in an ERGM by averaging the marginal effect over the set of eligible dyads $\mathcal{D}$ (Duxbury, 2021; Duxbury and Wertsching, 2023):

$$\text{AME}^x = \frac{1}{|\mathcal{D}|} \sum_{(i,j) \in \mathcal{D}} \text{ME}_{ij}^x \tag{5}$$

The set of eligible dyads consists of all pairs of actors (i, j) such that actor i could send a tie to actor j (Duxbury, 2021; Duxbury and Wertsching, 2023). Thus, in our setting, the set of eligible dyads consists of all possible pairs of students in the same class, since students could only nominate classmates that they would like to avoid in the CILS4EU questionnaire.

3.2.1 Marginal effects for continuous covariates

When x is a continuous covariate, the marginal effect of covariate x on dyad (i, j) is computed via a partial derivative (Mood, 2009; Duxbury, 2021; Duxbury and Wertsching, 2023):

$$\text{ME}_{ij}^x = \frac{\partial \widehat{\mathbb{P}}(Y_{ij} = 1 \mid \mathbf{X}, Y_{-ij})}{\partial x_{ij}} \tag{6}$$

Here, x_{ij} denotes the value of covariate x for dyad (i, j). Thus, the marginal effect for continuous covariates measures the change in the predicted conditional probability of there being a tie from i to j for a small change in x_{ij}. Leveraging Equation (3), we can rewrite Equation (6) as:

$$\text{ME}_{ij}^x = \widehat{\theta}_x \cdot \widehat{\mathbb{P}}(Y_{ij} = 1 \mid \mathbf{X}, Y_{-ij}) \cdot \widehat{\mathbb{P}}(Y_{ij} = 0 \mid \mathbf{X}, Y_{-ij}), \tag{7}$$

where $\widehat{\theta}_x$ denotes the coefficient estimate associated with covariate x (Mood, 2009; Duxbury, 2021; Perraillon and Hedeker, 2025).

3.2.2 Marginal effects for factor covariates

In their implementation of marginal effects for ERGMs, via the `ergMargins` R package (Duxbury, 2025), Duxbury (2021) and Duxbury and Wertsching (2023) use Equation (7) to calculate the marginal effects for both factor and continuous covariates. However, Equation (7) was derived by defining the marginal effect as a partial derivative. Yet, defining the marginal effect as a partial derivative does not make sense when x is a factor covariate since it is unclear what would constitute a "small change" in x_{ij}.

To this end, we implement marginal effects in ERGMs differently for factor covariates than is done in the `ergMargins` package. Our implementation aligns with how marginal effects are typically calculated for factor covariates in logistic regression models (Mood, 2009; Perraillon and Hedeker, 2025). Namely, when x is a factor covariate, we compute the marginal effect of x on dyad (i, j) by taking the difference in the predicted conditional probability of a tie from i to j when x_{ij} is at some level k versus at the baseline level k_0:

$$\mathrm{ME}_{ij}^{x} = \widehat{\mathbb{P}}(Y_{ij} = 1 \mid x_{ij} = k, Y_{-ij}) - \widehat{\mathbb{P}}(Y_{ij} = 1 \mid x_{ij} = k_0, Y_{-ij}) \qquad (8)$$

So, if x is a binary covariate, we take $k = 1$ and $k_0 = 0$. If x is a factor covariate with more than two levels, k and k_0 must be selected to calculate the marginal effect of x (Mood, 2009; Perraillon and Hedeker, 2025).

3.3 Model specification

Table 3 provides a summary of the parameters we include in our block diagonal ERGM, as well as their interpretation and graphical representation. We include friendship-related attributes to study how a student's avoidance tendencies relate to their position in the friendship network and endogenous network attributes to examine structural dependencies in classroom avoidance networks. The gender controls are included since Kros et al. (2021) find that students' gender influences avoidance tendencies.

In terms of friendship-related effects, we include both local and global network properties related to students' connectedness in the friendship network. First, we include whether the student considers the classmate a friend and whether the classmate considers the student a friend to quantify the overlap between avoidance and friendship ties. Additionally, we include the number of length 2 paths from the student to the classmate in the friendship network to examine if students tend to avoid friends of their friends (i.e., friends-of-friends). As previously mentioned, we measure students' connectedness in the friendship network by identifying weakly connected components of the classroom friendship network. Two classmates are weakly connected if there exists an undirected path (of any length) between them in the friendship network. In our model, we include an indicator of whether the student and classmate are in the same weakly connected component to examine whether students tend to avoid classmates they are weakly connected to. We also directly examine students' avoidance tendencies within the largest friend group of the classroom friendship network by including an indicator of whether each student is in the largest weakly connected component of the classroom friendship network in our model. Finally, we account for students' popularity and activity in the friendship network by including friendship indegree and outdegree parameters in our model.

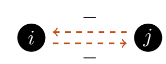
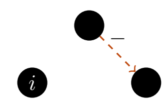
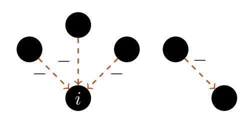
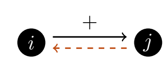
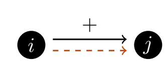
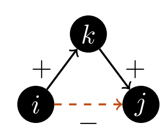
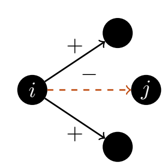
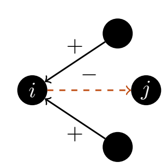
For endogenous network effects, we include reciprocity, the number of students with zero avoidance outdegree, a geometrically weighted indegree term, and an edges term, which functions as the model intercept. We include reciprocity to capture students' tendency to avoid classmates who avoid them. We account for heterogeneity in the number of classmates who avoid a student (see Figure 1B) by including a geometrically weighted indegree distribution term in our model. We estimate the decay hyperparameter for the geometrically weighted indegree term via a curved exponential family model (Snijders et al., 2006). The number of students with zero avoidance outdegree is included to improve model fit.

3.4 Model estimation

We fit our block diagonal ERGM with the `ergm` package in R (Hunter et al., 2008). Since the normalizing constant $\kappa(\theta, \mathcal{Y})$ is often computationally intractable to calculate in practice, the maximum likelihood estimate for θ is approximated via pseudo-maximum likelihood estimation (MPLE) or with a Markov Chain Monte Carlo (MCMC) algorithm (Hunter et al., 2008). Since we include dyad-dependent terms in our block diagonal ERGM, such as reciprocity and a geometrically weighted indegree term, the block diagonal ERGM coefficients are estimated with an MCMC algorithm (Hunter et al., 2008).

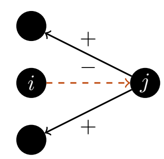
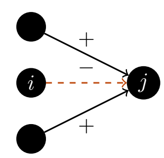
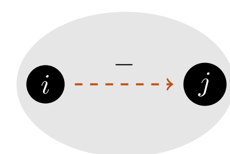
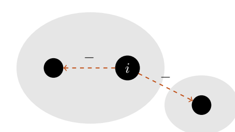
We impose two constraints on the sample space of possible networks to align with the CILS4EU questionnaire's constraints. First, we restrict our sample space of possible networks to those that are also block diagonal, since students could only nominate classmates they would like to avoid, not students in other classes. Second, we restrict all students' avoidance outdegrees to be at most five, since students could only nominate up to five classmates they would like to avoid in the CILS4EU questionnaire.

Table 3: Summary of block diagonal ERGM parameters, interpretations, and graphical representations.

Parameter	**Interpretation**	**Graphic**
Endogenous:		
Reciprocity	Tendency for avoidance ties to be reciprocated	
Outdegree is zero	Tendency of students to send no avoidance ties	
Geometrically weighted indegree	Tendency of incoming avoidance ties to concentrate among a small number of students	
Friendship:		
Student is friends with classmate	Tendency of students to avoid classmates they consider a friend	
Classmates is friends with student	Tendency of students to avoid classmates who consider them a friend	
Number of length 2 paths from student to classmate in friendship network	Tendency of students to avoid classmates who are friends with their friends (i.e., friends-of-friends)	
Student friendship outdegree	Tendency of students to avoid others based on number of outgoing friendship ties	
Student friendship indegree	Tendency of students to avoid others based on number of incoming friendship ties	

Continued on next page

Table 3 – continued from previous page

Parameter	**Interpretation**	**Graphic**
Classmate friendship outdegree	Tendency of classmates to be avoided based on number of outgoing friendship ties	
Classmate friendship indegree	Tendency of classmates to be avoided based on number of incoming friendship ties	
Student and classmate are weakly connected in friendship network	Tendency of students to avoid classmates in the same weakly connected component of the friendship network (i.e., same general friend group)	
Student in largest weakly connected component of class's friendship network	Tendency of students to avoid others based on if they are in largest connected component of their class's friendship network	

4 Results

Table 4 shows the results of fitting a block diagonal ERGM on the 193 classroom avoidance networks. We provide convergence and goodness-of-fit diagnostics in Appendix A.2. We discuss the friendship-related results in Section 4.1, the endogenous network results in Section 4.2, and the gender control results in Section 4.3. Given that our model is fit with 193 classrooms and 4,015 students, we have sufficient statistical power to detect even very subtle associations. To this end, we focus our results discussion on associations that are both practically meaningful and strongly statistically significant ($p < 0.001$). We also discuss null findings that had previously been found to be statistically significant.

4.1 Friendship-related effects

We find that avoidance networks are highly associated with the friendship networks. As expected, students tend *not* to avoid their friends. Student i being friends with classmate j is associated with an average decrease of 0.040 in the probability that student i avoids classmate j (AME $= -0.040$, $p < 0.001$). Similarly, students tend *not* to avoid classmates who consider them a friend. On average, classmate j being friends with student i is associated with a decrease of 0.036 in the probability that student i avoids classmate j (AME $= 0.036$, $p < 0.001$). We also find evidence that this effect is transitive in the sense that students are also *less* likely to avoid classmates who are friends-of-friends. Each additional length 2 path from the student i to classmate j in the friendship network is associated with an average decrease of 0.016 in the probability that student i avoids classmate j (AME $= 0.016$, $p < 0.001$). Our finding that students tend not to avoid friends-of-friends is consistent with balance theory (Cartwright and Harary, 1956). None of the friendship degree-related attributes are associated with more than a one percentage point change in the average probability that student i avoids classmate j; that is, $-0.01 \leq AME \leq 0.01$ for all friendship degree-related attributes in our model.

In terms of the attributes related to students' connectedness in the friendship network, we find that students tend *not* to avoid classmates in the same weakly connected component of the friendship network. On average, student i being in the same weakly connected component of the friendship network as classmate j is associated with a decrease of 0.016 in the probability that student i avoids classmate j (AME $= 0.016$, $p < 0.001$). Interestingly, we have a statistically significant positive coefficient and AME for the student being in the largest connected component of their class's friendship network. Parametrized differently, this means that students *not* in the largest weakly connected component of their class's friendship network are *less* likely to avoid others. One explanation for this is that students on the periphery of their class's friendship network already do not have that many friends. In turn, these students may not be in the business of actively avoiding people, as they would like to be liked by their peers.

4.2 Endogenous effects

Turning to endogenous network effects, we find that there are strong reciprocity and degree-related effects in classroom avoidance networks. Like Kros et al. (2021), our results indicate that avoidance tends to be reciprocated between classmates (AME = 0.077, $p < 0.001$). Further, many students do not avoid any of their classmates (AME = 0.049, $p < 0.001$). We have a very strong, statistically significant negative coefficient and AME on the geometrically weighted indegree term. This indicates that classroom avoidance networks are highly centralized (Snijders et al., 2006; Levy, 2016). Namely, there are more students who are not avoided by any of their classmates and more who are avoided by several of their classmates than expected by chance (Levy, 2016).

4.3 Gender effects

Finally, we do not find sufficient evidence that gender has a meaningful effect on students' avoidance tendencies. Unlike Kros et al. (2021), we do not find sufficient evidence that students avoid classmates of different genders more than those with the same gender (AME = -0.003, $p = 0.092$). While classmate j identifying as a boy is statistically significant ($p < 0.001$), it only increases the probability that student i avoids classmate j by 0.007, on average (AME = 0.007, $p < 0.001$), and thus is not very practically significant.

4.4 Average marginal effects

Per Table 4, our AME results align with those from our original block diagonal ERGM parameter estimates. All of our calculated AMEs and original ERGM parameter estimates have the same sign and roughly the same p-value magnitude. Since AMEs express results on the probability scale rather than the log-odds scale, we chose to report our AME results in the previous sections. Note that we compute the AME for each parameter by averaging the parameter's marginal effect over all dyads consisting of students in the same class. If we consider the distribution of the marginal effects for friendship-related attributes in our model, we find that the distributions are all skewed, suggesting that the included friendship-related attributes influence the predicted conditional probability that student i avoids classmate j to varying degrees (Figure A.5).

Table 4: Results of fitting a block diagonal ERGM on 193 classroom avoidance networks. Note that the ERGM coefficients are on the log-odds scale, and the Average marginal effects (AME) are on the probability scale. We do not estimate the AME for decay since it is a hyperparameter of the geometrically weighted indegree term.

	Coefficient			**AME**		
Covariate	Est.	SE	p	Est.	SE	p
Endogenous:						
Edges	0.471	0.058	<0.001	0.043	0.005	<0.001
Reciprocity	0.796	0.045	<0.001	0.077	0.005	<0.001
Outdegree is zero	0.713	0.055	<0.001	0.049	0.003	<0.001
Geometrically weighted indegree	-3.494	0.123	<0.001	-0.310	0.005	<0.001
Decay	2.252	0.066	<0.001	-	-	-
Friendship:						
Student is friends with classmate	-0.583	0.064	<0.001	-0.040	0.004	<0.001
Classmate is friends with student	-0.516	0.058	<0.001	-0.036	0.004	<0.001
Number of length 2 paths from student to classmate in friendship network	-0.181	0.020	<0.001	-0.016	0.002	<0.001
Student friendship outdegree	0.111	0.009	<0.001	0.010	0.001	<0.001
Classmate friendship outdegree	-0.009	0.006	0.114	-0.001	0.001	0.070
Student friendship indegree	-0.014	0.007	0.056	-0.001	0.001	0.098
Classmate friendship indegree	-0.055	0.006	<0.001	-0.005	0.001	<0.001
Student and classmate are weakly connected in friendship network	-0.186	0.028	<0.001	-0.016	0.002	<0.001
Student in largest connected component of class's friendship network	0.197	0.049	<0.001	0.016	0.004	<0.001

Continued on next page

Table 4: Results of fitting a block diagonal ERGM on 193 classroom avoidance networks. *(continued)*

	Coefficient			**AME**		
Covariate	Est.	SE	p	Est.	SE	p
Gender:						
Student identifies as a boy	0.060	0.025	0.016	0.005	0.002	0.023
Classmate identifies as a boy	0.089	0.015	<0.001	0.007	0.001	<0.001
Student and classmate are same gender	-0.038	0.021	0.070	-0.003	0.002	0.092

5 Discussion

In this project, we studied the structural dependencies in classroom avoidance networks, as well as how a student's tendency to avoid others and be avoided relates to their position in the friendship network. Using data from 193 secondary school classrooms collected during the first wave of the Children of Immigrants Longitudinal Survey in Four European Countries (CILS4EU), we fit a block diagonal exponential random graph model (ERGM) to examine how friendship and endogenous network attributes shape avoidance tie formation in classrooms. We find that classroom friendship networks are highly associated with classroom avoidance networks, and there are strong reciprocity and degree-related endogenous effects in classroom avoidance networks.

We also improved the existing implementation of average marginal effects (AMEs) in ERGMs by aligning the computation of marginal effects for factor covariates with the standard approach in logistic regression. Our computed AMEs and original ERGM parameter estimates have the same sign and magnitude of statistical significance. Thus, our AME results match those from our original ERGM parameter estimates. In contrast, Duxbury and Wertsching (2023) find that the AME and original block diagonal ERGM parameters in their empirical example differ in both their direction and statistical significance, which partially motivates their recommendation that AMEs should be reported in addition to the original parameter estimates for block diagonal ERGMs. However, in their simulation study, Duxbury and Wertsching (2023) find that their block diagonal ERGM parameters and AMEs all have the same sign. Additionally, only homophily effects (i.e., same race and same sex) in their simulation study produced Type II errors, and there were no Type I errors (Duxbury and Wertsching, 2023). Therefore, the discrepancy between our results and the empirical results from Duxbury and Wertsching (2023) could be explained by the fact that we do not include many homophily effects in our block diagonal ERGM, and also fit the model using several (193) classroom networks, which gives us high statistical power. Regardless, however, AMEs are helpful for practical interpretation because they express results on the probability scale rather than the log-odds scale. For this reason, we chose to report our AME results in the previous sections.

However, our analysis has a few limitations. For one, our results rely on the assumption that all 193 classroom networks share the same ERGM parameters. Yet, several of the network descriptives have a large standard deviation relative to their mean (Table 2), indicating that there is heterogeneity in the structure of classroom avoidance networks. In future work, we will account for this heterogeneity by using a Bayesian hierarchical ERGM, which allows each classroom to have its own ERGM parameters that deviate from the population-level parameters (Agneessens et al., 2024).

Moreover, because we use cross-sectional network data, we cannot distinguish between different mechanisms that are consistent with our findings. For example, our result that students tend not to avoid classmates who are friends-of-friends could be explained by triadic closure or balance theory, which suggests that students would be less likely to avoid friends-of-friends in order to maintain consistency in their social relationships. Alternatively, students may tend not to avoid classmates who are friends-of-friends because of homophily: students may share underlying similarities with their friends' friends and consequently would be more likely to interact with them than avoid them. Future work could use longitudinal network analysis to discern which mechanism best explains the associations we found in this study.

Finally, the generalizability of our findings may be limited by the use of classroom network data from Dutch secondary schools, where students are sorted into classrooms based on their academic level and interests. Avoidance structures could be quite different in classrooms with more diverse student compositions or in other organizational contexts. Therefore, it would be beneficial for future work to explicitly study the avoidance structure in such settings.

References

Agneessens, F., Trincado-Munoz, F. J., and Koskinen, J. (2024). Network formation in organizational settings: Exploring the importance of local social processes and team-level contextual variables in small groups using bayesian hierarchical ergms. *Social Networks*, 77:104–117.

Anderson, C. J., Wasserman, S., and Crouch, B. (1999). A p* primer: logit models for social networks. *Social Networks*, 21(1):37–66.

Candellone, E., Babul, S. A., Togay, O., Bovet, A., and Garcia-Bernardo, J. (2025). Negative ties highlight hidden extremes in social media polarization. *Network Science*, 13.

Cartwright, D. and Harary, F. (1956). Structural balance: a generalization of heider's theory. *Psychological Review*, 63(5):277–293.

Duxbury, S. (2025). *ergMargins: Process Analysis for Exponential Random Graph Models*. R package version 1.6.

Duxbury, S. W. (2021). The problem of scaling in exponential random graph models. *Sociological Methods & Research*, 52(2):764–802.

Duxbury, S. W. and Wertsching, J. (2023). Scaling bias in pooled exponential random graph models. *Social Networks*, 74:19–30.

Frank, O. and Strauss, D. (1986). Markov graphs. *Journal of the American Statistical Association*, 81(395):832–842.

González, M., Herrmann, H., Kertész, J., and Vicsek, T. (2007). Community structure and ethnic preferences in school friendship networks. *Physica A: Statistical Mechanics and its Applications*, 379(1):307–316.

Goodreau, S. M., Kitts, J. A., and Morris, M. (2009). Birds of a feather, or friend of a friend? using exponential random graph models to investigate adolescent social networks. *Demography*, 46(1):103–125.

Heider, F. (1946). Attitudes and cognitive organization. *The Journal of Psychology*, 21(1):107–112.

Huitsing, G., Snijders, T. A. B., Van Duijn, M. A. J., and Veenstra, R. (2014). Victims, bullies, and their defenders: A longitudinal study of the coevolution of positive and negative networks. *Development and Psychopathology*, 26(3):645–659.

Huitsing, G. and Veenstra, R. (2012). Bullying in classrooms: Participant roles from a social network perspective. *Aggressive Behavior*, 38(6):494–509.

Hunter, D. R., Handcock, M. S., Butts, C. T., Goodreau, S. M., and Morris, M. (2008). ergm: A package to fit, simulate and diagnose exponential-family models for networks. *Journal of Statistical Software*, 24(3):1–29.

Kalter, F., Heath, A. F., Hewstone, M., Jonsson, J. O., Kalmijn, M., Kogan, I., and van Tubergen, F. (2017). Children of immigrants longitudinal survey in four european countries (CILS4EU) - full version. data file for on-site use. GESIS, Cologne. (ZA5353; Version 3.3.0).

Kros, M., Jaspers, E., and van Zalk, M. (2021). Avoidance, antipathy, and aggression: A three-wave longitudinal network study on negative networks, status, and heteromisos. *Social Networks*, 64:122–133.

Labianca, G. and Brass, D. J. (2006). Exploring the social ledger: Negative relationships and negative asymmetry in social networks in organizations. *Academy of Management Review*, 31(3):596–614.

Levy, M. A. (2016). gwdegree: Improving interpretation of geometrically-weighted degree estimates in exponential random graph models. *The Journal of Open Source Software*, 1(3):36.

Lusher, D., Koskinen, J., and Robins, G., editors (2012). *Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications*. Cambridge University Press.

Mood, C. (2009). Logistic regression: Why we cannot do what we think we can do, and what we can do about it. *European Sociological Review*, 26(1):67–82.

Perraillon, M. and Hedeker, D. (2025). Health services research and program evaluation: Causal inference and estimation, online supplement. `https://www.perraillon.com/book.html`.

Snijders, T. A. B., Pattison, P. E., Robins, G. L., and Handcock, M. S. (2006). New specifications for exponential random graph models. *Sociological Methodology*, 36(1):99–153.

Toroslu, A. and Jaspers, E. (2022). Avoidance in action: Negative tie closure in balanced triads among pupils over time. *Social Networks*, 70:353–363.

Vega Yon, G. G., Slaughter, A., and de la Haye, K. (2021). Exponential random graph models for little networks. *Social Networks*, 64:225–238.

Wasserman, S. and Pattison, P. (1996). Logit models and logistic regressions for social networks: I. an introduction to markov graphs and p. *Psychometrika*, 61(3):401–425.

Wójcik, M. and Flak, W. (2019). Frenemy: A new addition to the bullying circle. *Journal of Interpersonal Violence*, 36(19–20):NP11131–NP11154.

A Appendix

A.1 Overlap in tie missingness

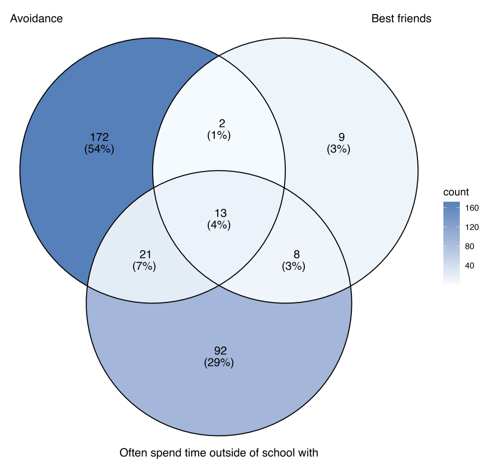


Figure A.1: Venn diagram showing the overlap in tie missingness in our data. Note that there is very little overlap between missingness in avoidance nominations and missingness in friendship nominations.

A.2 Model diagnostics

A.2.1 Convergence diagnostics

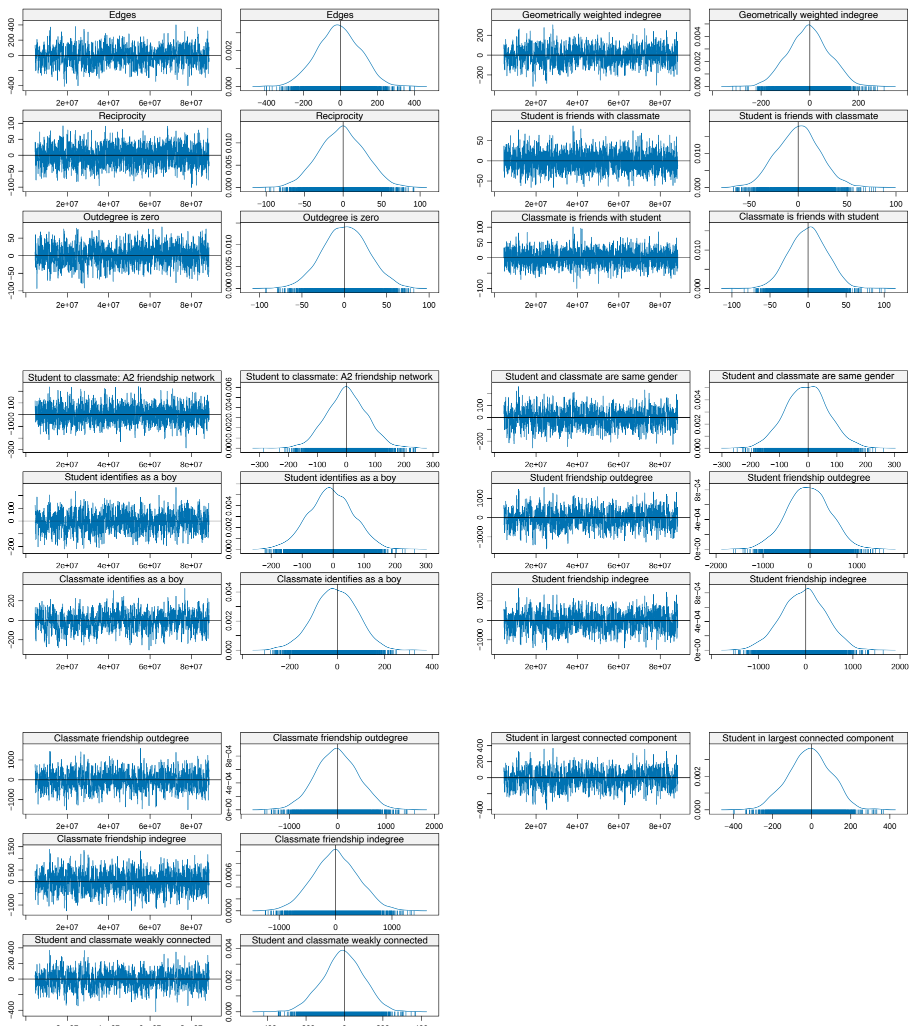


Figure A.2: MCMC diagnostics for each parameter included in our block diagonal ERGM. The trace plots indicate that all chains are well-mixed, and the posterior distributions appear stable, which is consistent with model convergence.

A.2.2 Goodness-of-fit diagnostics

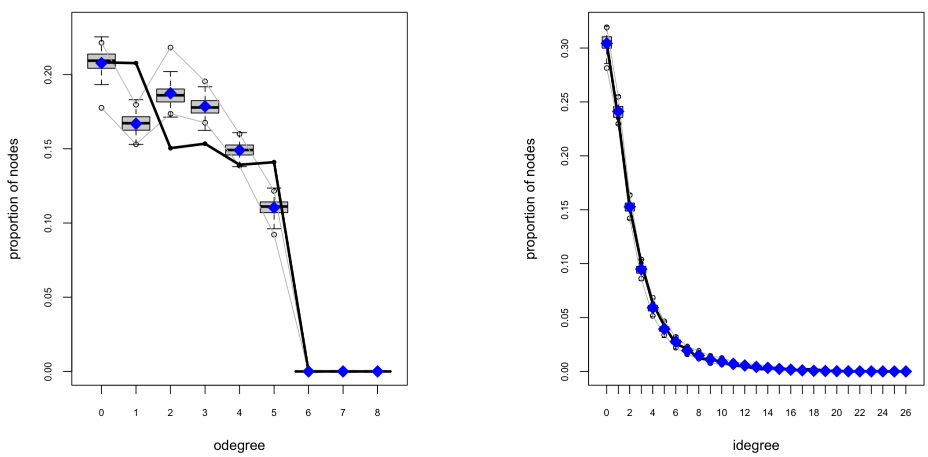


Figure A.3: Goodness-of-fit for the outdegree (left) and indegree (right) distributions. The black line shows the observed proportion of students with each outdegree/indegree. The boxplots summarize the distribution of these proportions across 100 networks simulated from our fitted block diagonal ERGM.

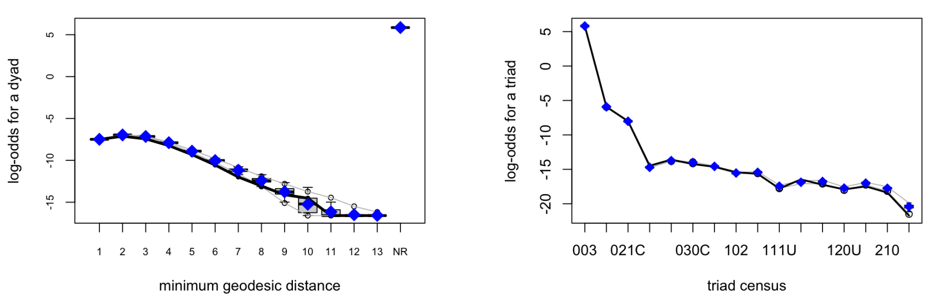


Figure A.4: Goodness-of-fit diagnostics for geodesic distance (left) and triad census (right) distributions. The black line shows the observed log-odds for a dyad/triad with each statistic. The boxplots summarize the distribution of these log-odds across 100 networks simulated from our fitted block diagonal ERGM.

A.3 Friendship marginal effect distributions

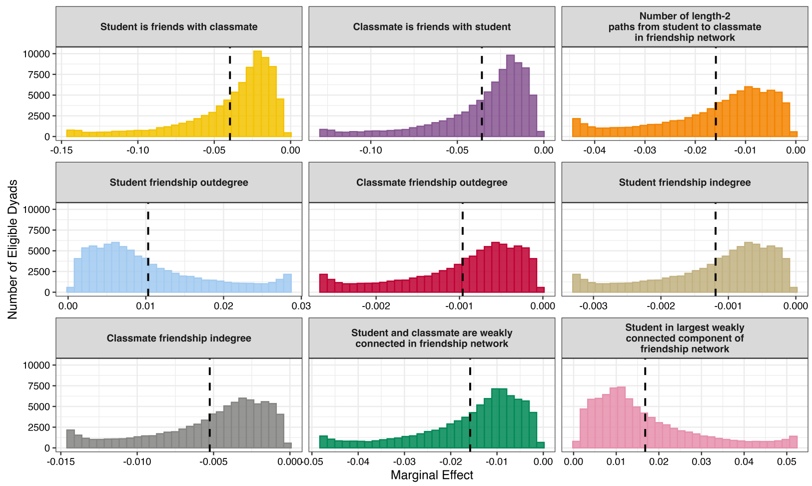


Figure A.5: Marginal effect distributions for all the friendship-related attributes included in our model. The dashed line denotes the AME for the given attribute. Note that the horizontal axis is on a different scale for each attribute.