Social Networks 76 (2024) 160–173

Contents lists available at ScienceDirect

Social Networks

journal homepage: www.elsevier.com/locate/socnet

How does socioeconomic homophily emerge? Testing for the contribution of different processes to socioeconomic segregation in adolescent friendships

Timothée Chabot

University Toulouse Jean Jaurès, LISST (UMR5193), France

ARTICLE INFO

Keywords:
Social network analysis
Homophily
Sociology of education
Socioeconomic origin
Stochastic actor-oriented models

ABSTRACT

Homophily – the fact that friendships happen at a higher rate among similar individuals – does not necessarily imply homophilic selection – the tendency to look for similar friends. This is particularly true for socioeconomic homophily: because individuals' social class impacts most aspects of their lives, there are several ways in which it can favor homogeneity in friendship networks. Applying this view to the relationships of French middle-school students, the present article tries to unravel the contribution of various relational processes to the emergence of socioeconomic homophily. Stochastic Actor-Oriented Models, a class of generative models designed for network panel data, are applied to the friendship networks of 820 students surveyed over a year and a half. Simulations derived from the estimated models are then used to assess the impact of different processes on aggregated levels of homophily. To that aim, a new metric is proposed that help researchers decompose an observed property of a network into a set of contributions from low-order processes, called "contribution scores". Results suggest that direct homophilic selection can be important in explaining socioeconomic homophily, but not in all cases. Indirect inducers, such as residential propinquity or ethnic selection, also play a significant role. Moreover, endogenous network processes – namely reciprocation and transitive closure – strongly contribute to homophily by reinforcing other homophily-inducing processes.

1. Introduction

Homophily – defined as the fact that relationships occur at a higher rate among similar than dissimilar individuals – is a pervasive characteristic of friendships and social relations in general. It has been consistently observed across a wide range of attributes, countries, contexts, and age groups (McPherson et al., 2001). Among these, researchers have long held a particular interest in homophilic patterns within children's and teenagers' networks. This is partly due to practical reasons (collecting friendship data is usually easier in the context of a school) but also comes from the particular significance of friendship ties at that age. Indeed, during teenage years, individuals' reference group in terms of normative behavior gradually shifts from the family to the peers, such that same-age friends play a crucial role in the socialization of adolescents and young adults (Coleman, 1961; Octobre et al., 2010). Therefore, homophily patterns determine who socializes with whom, with lasting implications over the life course and, consequently, on social structure.

In that regard, the three main attributes that sociological research has considered are gender (e.g., Shrum et al., 1988), ethnicity/race (Baerveldt et al., 2004; Moody, 2001), and academic performance (Flashman, 2012), each of those eliciting consistent homophily among youths. Social class, however – or, in the case of children and adolescents, socioeconomic origin – has received little attention in comparison (Malacarne, 2017).

This lack of studies is surprising because research on adult friendships found occupation, income, and educational level to impact personal networks strongly (McPherson et al., 2001). We also know that socioeconomic attributes shape individuals' values, beliefs, habits, or tastes (Bourdieu, 2016), which might determine the type of friends preferred. Moreover, among school students, ethnicity and academic performance are correlated to socioeconomic origin (Ichou, 2018), so identifying homophilic behaviors on either of these variables demands a careful examination of the effect of the other two. Finally, there is widespread concern for socioeconomic segregation across schools (Oberti et al., 2012), as the composition of schools seems to impact students' performance net of individual characteristics (Coleman, 1995). For this reason, socioeconomic mixing is commonly seen as

E-mail address: timothee.chabot@univ-tlse2.fr.

https://doi.org/10.1016/j.socnet.2023.09.002

Available online 26 September 2023

desirable. Nevertheless, one should not only consider the formal composition of the schools' population but also, given this composition, the actual level of interactions between socioeconomic groups.

The first aim of the present article is thus to investigate socioeconomic homophily in teenagers' friendship networks, using data from four French middle schools. I will document the descriptive levels of homophily observed in these schools, then attempt to explain how this homophily emerged by assessing the role of different relational mechanisms, such as the direct selection of similar friends, opportunity structures, or network endogenous processes.

A secondary aim is methodological. Network researchers are used to specify generative models in order to reconstitute the data-generating process that most plausibly underlie an observed network. This is the rationale of the Exponential Random-Graph Model (ERGM) or, for panel data, of the Stochastic Actor-Oriented Model (SAOM). This generative approach is necessary to test for competing hypotheses about the role of different low-order mechanisms in the emergence of network structure – for instance, whether homophily is born out of selection or of constrained opportunities. Nevertheless, this raises a number of methodological challenges, notably because the intrinsic dependency of network observations makes it difficult to relate a particular process (e.g. selection) to a particular outcome (e.g. homophily).

One promising avenue of research in this direction is to use the SAOM as an empirically-calibrated Agent-Based Model (ABM), and to explore hypothetical scenarios where low-order mechanisms vary through simulations. However the technical details of this type of simulation analysis are not straightforward, and the literature is still debating what the best approach is (Block, 2018; Steglich and Snijders, 2022). The present article makes a new proposal in that direction, which I call "contribution scores", and applies it to the empirical question of the emergence of socioeconomic homophily among teenagers.

2. Literature and theory

2.1. Socioeconomic homophily in the literature

As mentioned, few network studies look at socioeconomic homophily among school students. In several cases, socioeconomic origin is used as a control variable, as researchers are primarily interested in ethnicity or race. Going through the result tables of these articles, the statistical effects associated with socioeconomic homophily generally appear to be small. Smith et al. (2014) found no socioeconomic homophily while controlling for ethnicity, gender, and sociometric popularity (German, Dutch, and Swedish high-school students). In Kruse et al.'s models (2016), socioeconomic homophily, although statistically significant, has a much smaller effect size than ethnic homophily (German and Dutch high-school students). Similarly, McFarland et al. (2014), using the extensive Add Health data set, found consistent socioeconomic homophily across US secondary schools but of a magnitude largely inferior to that of racial homophily.

However, estimates come from a multivariate model in all these cases. Since researchers are not primarily interested in socioeconomic origin, they do not present descriptive estimates for gross levels of socioeconomic homophily, which means related effects in the multivariate models could have explained it away. Moreover, the socioeconomic composition of the different schools – which is likely to have a decisive impact on homophilic patterns – is unknown. Finally, the lack of interest in this research question may mean that the collection and coding of the variables used to measure socioeconomic origin – usually parental occupation or education – is not as precise and extensive as it could be. For these reasons, these results may not be a good indication of the magnitude of socioeconomic homophily in school contexts.

2.2. Inducing processes of socioeconomic homophily

The term "homophily" can have different meanings in the literature. It is sometimes used to refer to an over-representation of ties among similar individuals relative to a random reference distribution (e.g., Goodreau et al., 2009), and sometimes to designate individuals' propensity to bond with similar others (e.g., Wimmer and Lewis, 2010). By convention, I will use *homophily* in the first sense, to refer to the descriptive network-level outcome; and *homophilic selection* for the relational process by which individuals look for similar others. This distinction is crucial for our research question because socioeconomic homophily does not imply socioeconomic homophilic selection (hereafter, "socioeconomic selection"). Indeed, several processes may explain how socioeconomic homophily emerges in adolescent friendship networks.

First, individuals' socioeconomic attributes impact their meeting opportunities; that is, the likelihood of coming into contact with specific others. In the literature, various concepts have been proposed to capture this idea, notably the notions of "propinquity" (i.e., spatial proximity among individuals; Kruse et al., 2016) and "social foci" (i.e., social contexts where relationships can form; Feld, 1981). In any case, friendship formation should occur more often as the frequency of contact among individuals increases, though most likely with decreasing marginal returns (Block, 2018). Regarding socioeconomic origin, residential segregation and differential school choice by parents result in children and teenagers being more likely to meet peers from similar origins at school and in the neighborhood (Oberti et al., 2012).[1] Even within a given school, socioeconomic origin affects students' results, motivation, and orientation, resulting in socioeconomically segregated tracks or ability groups (Duru-Bellat and Mingat, 1997; Palheta, 2012). The type of sport, art, and extracurricular activities preferred by children and teenagers also markedly varies with socioeconomic origin (Octobre et al., 2010). Moreover, parents have substantial control over their children's activities and may push them toward socially segregated ones. Upper-class parents tend to regulate the activities their children attend, partly to avoid an excessive level of socioeconomic and racial diversity which they see as detrimental to their success or well-being (Ball et al., 2004; van Zanten, 2015). Finally, inter-generational closure – that is, the fact that friendships between parents facilitate contact among children (Coleman and Hoffer, 1987) – should also entail higher bonding opportunities among same-background children, given the socioeconomic segregation of adult networks.

Second, youths may have a subjective preference for similar peers based on attributes that, although distinct from socioeconomic origin, are correlated to it, thus inducing socioeconomic homophily through compositional effects. We have much evidence for homophilic selection based on ethnicity (Baerveldt et al., 2004) as well as academic results (Flashman, 2012; Smirnov and Thurner, 2017), both of which strongly correlate with socioeconomic origin (Ichou, 2018). Furthermore, the same line of reasoning applies to more subtle attributes, hard to objectify through quantitative measures but very much linked to one's socioeconomic origin: tastes, beliefs, habits, values, ways of feeling, thinking, and behaving (Bourdieu, 2016). For example, leisure activities do not

[1] References in this section refer to different countries. The studies about students' results, school choice, spatial segregation, and social inequalities mostly pertain to French students since my study takes place in this country. However, social network studies are based on American, Dutch, or German data due to a lack of such studies in France. I expect the generic mechanisms discussed here to be similar across Western countries, although this would deserve a dedicated investigation.

only impact friendship formation through meeting opportunities but also homophilic selection, as friendships should be more easily formed among students that share an interest in, say, a given sport or musical style.

In fact, to the extent that some of these attributes are more or less direct expressions of socioeconomic positions, they could be described as mediators of socioeconomic selection rather than mere compositional inducers.[2] However, this is still different from “raw”, direct discrimination based on wealth or parental occupation, which may also occur. For instance, Eder (1985) observed a strong stigma associated with poverty in a US secondary school, with wealthy students openly mocking “grits”.

Finally, dependencies between friendship ties create a non-linear relationship between the processes occurring at the dyad level and the amount of homophily found in the network. These higher-order processes are referred to as *endogenous network processes*, with reciprocation (looking for reciprocal friendships) and transitive closure (the principle by which friends of friends tend to become friends) being the ones most consistently found in friendship networks (Block, 2018). In general, these should aggravate homophily: they tend to generate clustered network structures, which, to the extent that friendship groups are already homophilic because of other processes, can make it even more difficult to bond with dissimilar others. For example, controlling for transitivity has been found to markedly reduce the dyad-level estimates for ethnic homophily among US secondary-school students (Goodreau et al., 2009). Theoretical works based on stylized simulations also found that transitive closure can result in high levels of homophily among nodes that have but a slight propensity to select similar partners (Asikainen et al., 2020; Foster et al., 2011).

2.3. Methodological challenges in identifying contributions to homophily

As we saw, many processes may generate homophily among school students. To separate explanatory factors and assess their respective importance, it is necessary to go beyond descriptive statistics about the aggregate state of friendship networks and resort to inferential modeling to reconstitute the underlying generative process. This same issue has been raised in the literature on racial homophily. For instance, Goodreau et al. (2009) and Wimmer and Lewis (2010) have used inferential network models (in their case, ERGMs) to disentangle various processes susceptible to eliciting racial homophily in friendship networks. However, a limitation of these works is that they lack a reliable way of measuring the importance of a given mechanism in terms of its impact on homophily specifically.

Indeed, models such as ERGM or SAOM assess the importance of different relational processes *regarding the entire network structure*, but not regarding a specific feature like racial or socioeconomic homophily. The works I just mentioned get around this issue by comparing the size of estimated coefficients across nested models (they consider the reduction in the size of the racial homophily parameter as more effects get added to the model), but this is not entirely satisfactory: the relative importance of the different effects depends on the order in which they are added to the model specification, and, most importantly, inter-model comparison of log-odd coefficients is at risk of being biased (Mood, 2010; note that both ERGM and SAOM have a logit model at their core). For this reason, I propose another approach to the issue of linking processes to a given network feature, which is presented in detail below.

Altogether, disentangling the inducing processes of socioeconomic homophily demands a three-step analysis. First, we ought to assess the “raw” quantity of socioeconomic homophily in students’ friendship networks, regardless of how it emerged; then, we need to reconstitute the set of tie-formation processes that has (most likely) generated these networks; finally, we need to assess how each mechanism relates to socioeconomic homophily specifically.

3. Methods

3.1. Data

Data comes from four French middle schools, two located in the Parisian agglomeration and two in the countryside in Savoie, a region South-East of France. By convention, I will refer to them as Paris 1, Paris 2, Savoie 1, and Savoie 2. These schools are all socioeconomically diverse, welcoming students from lower-class, middle-class, and upper-class origins. Paris 1 and Savoie 1 are public schools (i.e., State schools), whereas Paris 2 and Savoie 2 are private catholic schools (though not all students are catholic themselves). In each school, I consider the entire cohort of students that entered secondary education in Fall 2017 ($n = 820$, respectively 96, 288, 203, and 233 in Paris 1, Paris 2, Savoie 1, and Savoie 2). These students were surveyed by questionnaire four times between their first and third year of middle school, with a 6-month gap between each wave (age in wave 1 = 11/12; age in wave 4 = 12/13). All members of the sample are therefore in the same grade-level, but the entire cohort moves up the levels over the year and a half of the study. See Appendices 1 and 2 for details.

The questionnaires asked students about their gender, former primary school, place of residence, parental occupation, and the foreign languages they could speak or read. Pairwise walking distances between students’ homes were computed using Open Route Service (Oles, 2021), a cartographic tool based on Open Street Map data. Students’ ethnic background was coded by combining information about the foreign languages spoken together with the first and last names, resulting in 5 broad ethnic categories (Appendix 3). Additionally, grades (trimester averages) were obtained from the school administrations and standardized per school and wave.

For socioeconomic origin, the questionnaires asked students about their parents’ occupation (including step-parents or other guardians). To improve the measure’s reliability, a researcher asked those whose answer was missing or ambiguous for further details orally at a later wave. Answers were then converted into an ISEI score (Ganzeboom et al., 1992) for each parent. The ISEI score of a student is defined as the average score of the household parents. The resulting measure ranges from 14.64 to 88.70, with mean 44.98 and standard deviation 19.36 (theoretical extremums for ISEI scores are 10 and 90). It is then standardized across all schools for a more straightforward interpretation of models’ coefficients.

Finally, students were presented with an exhaustive list of all students in their grade level and could indicate for each one whether they considered them to be a “very good friend”, “friend”, “someone that [they] dislike”, or neither. “Friend” nominations were used mainly by students to refer to weak ties, as shown by the high average number of nominations (21.0 per student across schools and waves, against 12.6 for very good friends). This raises issues of interpretation (how significant are these relationships?) and, more practically, of modeling (it is harder to find a converged and well-fitted model for denser networks). Therefore, in the rest of the article, I focus exclusively on "very good friend" nominations, though I will speak of “friendship networks” for simplicity. These networks are observed four times, with a 6-months gap between each observation. Descriptives are in Appendix 4.

[2] This depends on the kind of causal relationship that one is willing to assume between socioeconomic origin and other variables. For example, it seems reasonable to consider ethnicity as a different dimension of social stratification entirely, which, although correlated to social class, is not caused by it; in which case, ethnic selection simply entails socioeconomic homophily. On the contrary, one’s passion for jazz music is most likely caused by their socioeconomic origin, in which case we would say that jazz-based selection *mediates* socioeconomic selection.

3.2. Network models

3.2.1. Permutation tests

Our first research question – whether there is socioeconomic homophily among students – is a descriptive one. The statistic used to measure socioeconomic homophily is the absolute difference in ISEI scores among the two nodes forming a tie, averaged over all ties in the network. Since this value depends in part on the structural characteristics of the networks, I compare the observed value to a simulated distribution obtained by randomly permuting the nodes' labels in the observed networks. This preserves the structure of ties but breaks its relationship with the statistic of interest, thus providing a null distribution to assess the statistical significance of the observed value.

The permutations are done for all four longitudinal observations of friendship networks at once, in order to preserve the structure of the evolution of the networks over time. Therefore, if student A is permuted with student B in wave 1, they are also permuted with B in all subsequent waves. The statistic indicative of socioeconomic homophily is then computed for each wave and averaged over the four waves. Empirical p-values are computed as the proportion of permuted networks for which this averaged statistic is smaller than in the observed networks (because of how the homophily statistic is defined, smaller values mean more homophily). These analyses are performed using the *statnet* package (Handcock et al., 2016) in R (R Core Team, 2020).

3.2.2. The stochastic actor-oriented model

To infer the relational processes underlying the observed networks, I resort to the *Stochastic Actor-Oriented Model* (SAOM), a class of statistical models designed for network panel data (Snijders, 2001, 2017). SAOM models the evolution of a network from the first observation onward, which is taken as granted. The subsequent observations are assumed to be the outcome of a continuous-time Markov process that can be decomposed into a series of *mini-steps* in-between actual network observations. At each mini-step, an actor can "decide" either to form one tie, drop one tie, or do nothing, depending on their network surroundings. A *rate function* determines the frequency of these decisions, and an *evaluation function* (also known as the o*bjective function*) determines their content (based on a conditional logit multinomial model). Both functions contain several parameters that need to be estimated from the data. Importantly for our purpose, the parameters of the evaluation function capture the change in the chances of forming or dropping a tie when this tie contributes to forming a particular configuration – for instance, if that tie closes a transitive triplet (transitivity effect) or if it connects two actors with the same gender (gender homophily effect).[3]

The estimation procedure used here is the Method of Moments (Ripley and Snijders, 2020). Under this procedure, the network is characterized by macro-level *target statistics*, each of which is a sum of all occurrences of a micro-level tie configuration in the network. Each statistic corresponds to a single parameter to be estimated, and vice-versa: for instance, adding a reciprocity effect to the model would mean adding both the target statistic of the number of reciprocated ties in the network, and a parameter for tie reciprocation in the evaluation function. The estimation algorithm then looks for the value of the parameters that will generate networks whose statistics are deemed similar enough to the observed ones according to some optimization criterion.

3.2.3. SAOM specification

The estimated models include effects that pertain to individual attributes and structural network effects. The specification is mostly the same for the four schools, with minor differences to account for some particularities of their network structures.

Regarding individual attributes, three types of effects are used: similarity, ego, and alter. Similarity effects capture the prevalence of ties among pairs of nodes with similar attribute values. Ego and alter effects capture the prevalence of ties that are either emitted (ego) or received (alter) by nodes with a given value of the attribute (note that these effects are defined differently for continuous and categorical attributes; see the RSiena manual). Similarity, emission, and reception are added to the model for ISEI, gender, ethnicity, and academic results. Furthermore, similarity effects are added for students' former primary school, current classroom, and classroom of the previous wave. Ego and alter effects for the current classroom are added in one school (Paris 1) but not the other three because it did not improve the model's fit. For the same reason, there is no ego nor alter effect for primary schools.

Moreover, in two schools (Paris 1 and Savoie 1), a small group of students (about 15) attend a special track. In Paris 1, it is an elitist track (a French-English bilingual section) integrated within a regular classroom (students attend most of their courses together, except for a few special bilingual courses). In Savoie 1, it is an entirely separate classroom for students with high learning difficulties. The track in Paris 1 mainly includes students with an upper-class background, whereas the one in Savoie 1 primarily includes students with a working-class background. In both cases, three dedicated effects are added to the model: two separated similarity effects (similarity of the special track and similarity of the majority group), plus a reception effect capturing the asymmetry in nominations emitted between the special and regular tracks.

Residential distance is a dyadic covariate giving the pairwise walking distance between students' homes. Two effects are included: one giving the walking time in hours, and the other the squared walking time, to account for the decreasing marginal relevance of residential distance (when students live far away from one another, they need to use public transportation or car, which compresses distances compared to walking).

Three types of structural effects are included. These are primarily meant to represent known endogenous network processes in friendship networks, but they can also offer some protection against omitted variables, such as unobserved attributes or meeting opportunities, that might influence the network structure (Block, 2018; Ripley and Snijders, 2020).

First, a reciprocity effect represents nodes' propensity to reciprocate nominations (forming i→j when i←j exists). Second, three effects pertain to triadic closure: two transitivity effects and an interaction effect between reciprocity and transitivity. Transitivity corresponds to the idea that "the friends of my friends are my friends": the two effects capture the chances of forming a tie i→j when it is included in two different types of transitive triplets (i→k→j, known as a "Forward-Forward" triplet, and i→k←j, or "Forward-Back" triplet). The interaction effect accounts for the fact that ties embedded in transitive triplets tend to be less reciprocal than those who are not, a common feature in friendship networks (see Block, 2015). Note that only one of the two transitivity effects (Forward-Forward) is interacted with reciprocity, for parsimony. All three of these effects follow the definition of geometrically-weighted edgewise shared partners ('gwesp') effects (Hunter, 2007; Ripley and Snijders, 2020). Third, three effects capture the dynamics of degrees: one for the in-degree distribution, another for the out-degree distribution, and one for the interaction of both (respectively "in-degree popularity", "out-degree activity", and "out-degree popularity"). These terms model the propensity of well-connected nodes to emit or receive more ties.

Finally, in certain schools, time dummies are added on a few effects to account for time heterogeneity, following the approach of Lospinoso et al. (2011). See Appendix 6 for detail.

[3] More precisely, these parameters are non-standardized contributions to log-probabilities within the evaluation function's multinomial model, which predicts actors' decisions at a single mini-step. See the RSiena manual (Ripley and Snijders, 2020), section 13.

3.3. Investigating the micro-macro link through simulations

3.3.1. Model analysis in an ABM framework

The overall philosophy behind SAOM is to find the combination of micro-level processes that best account for the macro-level structure of the network. These processes are usually meant to reflect actors' behaviors, not just to describe the over- or under-representation of various low-order configurations (Block et al., 2018 refer to this as the "micro-level interpretation" of parameters). The estimated coefficients thus capture the strength of these micro-level processes and, by extension, their importance in explaining the overall structure of the network.

However, they do not account for their importance regarding a particular network feature – i.e., a particular statistic. As such, SAOM coefficients do not tell us how much, say, transitive closure, ethnic homophilic selection, or classroom propinquity, account for socioeconomic homophily specifically. Indeed, by construction, the macro-level features of the network cannot be expressed as a sum of linear contributions from each model parameter. Instead, all parameters *jointly* generate network features, through complex and usually intractable interactions throughout the simulation procedure (which is necessary to account for the intrinsic dependency of network observations).

This type of issue is widespread in another field of research: agent-based modeling. Agent-based Models (ABM) are computational simulation models that generate aggregate features from the interactions of lower-level entities, usually to provide a stylized representation of a real-world process. These models frequently suffer from a "black box" problem in that it is difficult to understand how and why they generate a given outcome. Researchers thus need to perform a *model analysis*, defined as "*the set of strategies that can be used to understand (and describe) the set of events, behaviors, and feedbacks […] triggered by the mechanisms coded in the ABM*" (Manzo, 2022, p.66). One prominent method is to re-run the ABM while changing specific low-level processes to see how the simulated outcome reacts (*ibid*, p.67).

Due to its simulation-based estimation and actor-oriented nature, SAOM can be seen as an ABM, although it also integrates elements of general linear modeling and is more constrained by data than ABMs usually are (Snijders and Steglich, 2015; Stadtfeld and Amati, 2021). Therefore, the logic of model analysis can be extended to an already estimated SAOM. By manipulating the value of specific parameters and seeing how the simulated network structure evolves, it is possible to gain information about the processes enacted by the model.

A first option is to incrementally increase or decrease a parameter's value to obtain a somewhat continuous function of its relationship to the statistic of interest (e.g., Lakon et al., 2015). However, this requires many simulation rounds, so the computational burden might become too much if one wishes to repeat the operation for several parameters. Instead, one can directly set the parameter of interest to 0, which provides a clear reference point from which to compare several parameters. This is equivalent to simulating a counterfactual scenario where everything remains the same except for one relational process that is assumed to be absent.

Therefore, I propose to define the *contribution* of a SAOM parameter (or set of parameters) to a given statistic as the simulated change in this statistic when the parameter of interest is set to 0. Several authors applied this general idea before but with some variation in its detailed implementation (e.g. Block, 2018; Kruse et al., 2016; Snijders & Steglich 2022). I will first present my approach, which I call *contribution scores*, then discuss how it places relative to previous works (Section 3.3.4.).

3.3.2. Introducing contribution scores

Contribution scores are a metric to compare the change in a macro-level statistic when altering different model parameters. They are computed in three steps: first, define a network statistic that is indicative of the feature of interest (here socioeconomic homophily); then, simulate fictive networks from an altered model in which the parameter of interest is set to 0; finally, compare the average value of the statistic in the simulated networks to its value in the observed network. I detail these steps below, taking the example of socioeconomic homophily; however, this can generalize to any network statistic.

First, we need to quantify the raw quantity of socioeconomic homophily in a network. For a network x, let$S_{isei}(x)$ be the sum over all ties of their ISEI similarity score (i.e., the absolute difference in ISEI values among the two nodes, re-scaled on (0,1) and subtracted from 1). I call $x^{obs}$the true network observed in a given school. The ISEI similarity statistic for this network is then defined as $S^{obs} = S_{isei}(x^{obs})$.

Importantly, S^{obs} does not measure homophily strictly speaking: it gives the observed value of ISEI similarity among ties but does not tell whether this value is larger or smaller than expected by chance. Therefore, homophily is best defined as the difference between the observed value S^{obs} and its expected value under a null reference distribution. The null I use is a simple SAOM that only considers network density and rate parameters from each wave to the next (note that this null has the advantage of accounting for any homophily that may "mechanically" carry over from the first wave just because there has not been enough time to re-evaluate all ties before wave 2). It is estimated separately, then networks are simulated from this model, and the value of the ISEI similarity statistic is averaged over these fictive networks. LetX_{null} be the set of simulated networks and N the number of these networks. The predicted value of ISEI similarity under the null is then defined as:

$$S^{null} = \frac{1}{N}\sum_{x \in X_{null}} S_{isei}(x) \tag{1}$$

The homophily in the observed network can thus be measured by$S^{obs} - S^{null}$.

Second, we need to define the homophily that is expected when the parameter of interest is set to 0. Consider a parameter vector θ with n estimated parameters $\theta = (\theta_1, \theta_2 \dots \theta_n)$ (estimated from the empirical network, given a SAOM specification). Setting the kth parameter to 0 results in a new parameter vector θ^{-k}. Similar to what was done for the null model, we can simulate a large number of networks from the altered model defined by θ^{-k}, then average the ISEI statistic over these fictive networks. Let X_{sim} be the set of networks simulated from θ^{-k} and N the number of these networks. The predicted value of the ISEI similarity statistic according to the altered model is then defined as:

$$S^{-k} = \frac{1}{N}\sum_{x \in X_{\text{sim}}} S_{isei}(x) \tag{2}$$

A problem at this stage is that S^{-k} and S^{obs} will likely be incomparable. Indeed, $S_{isei}(x)$ is highly sensitive to the number of ties in the network, and the networks in X_{sim} will likely have a different density than x^{obs} (due to the alteration of the parameter vector). To circumvent this issue, I follow Block's (2018) recommendation to adjust the density (i.e., out-degree) parameter in the parameter vector θ^{-k}, such that the predicted density matches that of the observed network. In practice, this is done by re-estimating a SAOM with all parameters offset to their value in the original model, except the one whose contribution is tested, which is offset to 0, and the out-degree effect(s), which is (are) not offset. Therefore, the estimation will adjust the out-degree parameter(s) to reach the correct predicted density.

Then, as for the observed network, homophily is defined as the difference between the actual value of $S_{isei}(x)$, and its expected value under the null. The average socioeconomic homophily in the simulated networks can thus be measured by $S^{-k} - S^{null}$.

Finally, the difference in homophily between the observed network and the alternative scenarios can be assessed. To obtain a metric that is both easy to interpret and comparable across different networks and models, it is expressed as a proportion of the homophily in the observed network. The *contribution score* of any parameter θ_k to socioeconomic homophily is thus defined as:

$$C_k^{\theta} = \frac{(S^{-k} - S^{null}) - (S^{obs} - S^{null})}{S^{obs} - S^{null}} \quad (3)$$

The expression further simplifies as follows:

$$C_k^{\theta} = \frac{S^{-k} - S^{obs}}{S^{obs} - S^{null}} \quad (4)$$

This can be seen as the predicted proportion change in socioeconomic homophily when the parameter of interest is set to 0. Multiplied by 100, it can thus be interpreted as a percentage change. Negative scores indicate positive contributions to homophily (removing the process reduces homophily).

Finally, note that even though the scores are expressed in proportions, they will typically not sum up to − 1 over all model parameters. Indeed, setting a parameter to 0 means removing the processes by which it interacts with other modeled effects – and these interactions may impact socioeconomic homophily. Consequently, the joint contribution of two or more parameters cannot be assessed by summing their respective scores; instead, it is necessary to set them to 0 simultaneously and compute the corresponding score.

3.3.3. Interpretation: empirical relevance of model analysis

Contribution scores are a model analysis tool that helps understand how an estimated SAOM relies on different parameters to reach its optimization goal on a target statistic. Therefore, they are primarily informative about the model's artificial data-generating process (DGP). Only to the extent that this DGP reasonably represents real-world processes can the scores have empirical relevance. Note, however, that this is not specific to contribution scores: if real-world processes are entirely unrelated to simulated ones, then interpreting the parameters as indicative of actors' behaviors – as researchers typically do with SAOM – is also wrong. Another way to say this is that the scores explore the deductive implications of an estimated model.

Whether the model's DGP is realistic enough is never certain, but SAOM offers more guarantees in that regard than most ABMs. Manzo (2022, pp. 52–55) proposes three dimensions along which to evaluate the realism of an ABM: input realism (using empirical data to calibrate the model parameters), output realism (comparing the simulated outcome to empirical data), and theoretical realism (grounding the micro-level processes specified in the model in qualitative evidence, such as psychological rules of behavior). A well-specified SAOM can score high on these three dimensions through the estimation of the parameters from observed data (input realism), goodness-of-fit assessment (output realism), and choosing effects that represent known relational processes in network theory (theoretical realism). Under these conditions, it is rational to believe that not only the estimated parameters but also the generative process enacted within the model hold some relevant information about the real world (*ibid*, p.58).

Having said that, it is essential to understand what this information pertains to exactly. Contribution scores are informative about the model DGP and, by extension, about the alleged DGP that generated the observed network. In that regard, the "counterfactual" manipulation of setting a parameter to 0 is instrumental to the assessment of the role that *has been* played by various processes in the evolution of the observed network. This is a different question from knowing what *would* happen under changing circumstances. Surely, the two are not entirely unrelated: if one process has been essential in generating a network feature, then it may be reasonable to expect that affecting this process would change the feature substantially. However, this is not guaranteed, so further evidence would be needed to make the argument convincing. The reason is that the assumption that all processes remain identical apart from the manipulated one will often prove unrealistic in empirical contexts. For example, if one were to increase the socioeconomic diversity of meeting contexts in a school, students might change their selective behaviors in reaction. Therefore, the scores are not predictions of the effect of an intervention.

3.3.4. Differences with other simulation-based approaches

The idea of altering an estimated model and performing simulations to investigate a macro feature is not new in Social Network Analysis. Examples that rely on either SAOM or ERGM are Block (2018); Fujimoto et al. (2018); Kruse et al. (2016); Lakon et al. (2015); Snijders and Steglich (2015); Stadtfeld and Amati (2021); Steglich et al. (2010); Steglich and Snijders (2022). Going through these works, however, the detailed implementation of this idea differs. It seems natural that there is no consensus yet on this matter because the use of SAOM and ERGM as empirically-calibrated simulation models and the connection to the ABM literature are relatively recent. The approach proposed here is meant to contribute to this ongoing effort, as I believe more work is needed to figure out the best approach. Therefore, although the aim of this article is primarily empirical, I will take a moment to discuss some pending methodological issues.

One key question is whether model parameters other than the altered one(s) should be kept at the same value in the altered models (as is the case here) or whether they should be re-estimated. Block (2018, p.425) advocates that the parameters from the full model are closer to the actors' true behavior than re-estimated parameters. Indeed, in a nested model approach, re-estimated parameters take into account less information about the network than the original ones, making them less realistic.[4] Lakon et al. (2015) and Kruse et al. (2016) adopt similar approaches (with ERGMs for the latter). On the contrary, Snijders and Steglich (2015) worry that affecting a single parameter "*may lead to quite unrealistic networks, having implausible values for the average degrees, degree variances, and so on*" (p.33). Therefore, instead of setting a parameter to 0 while leaving others unchanged, they propose to re-estimate a new SAOM while taking the effect of interest out of the specification, following the logic of nested models (see also Fujimoto et al., 2018; Steglich et al., 2010; Steglich and Snijders, 2022).

Interestingly, both arguments revolve around the notion of "realism", but they envision it at a different level: Block worries about the realism of low-order processes (i.e., actors' behavior), whereas Snijders & Steglich worry about the realism of the aggregate network structure ("*average degrees, degree variance and so on*"). The type of simulated scenario implemented in each case reflects this: either the values of the parameters are fixed (i.e., not re-estimated), or the values of the target statistics are fixed (since re-estimating a parameter means forcing the simulated networks to reach the observed value of its corresponding statistic on average).

In the first case, we compare two DGPs that only differ by one low-order process: the parameter set to 0. That is why differences in the

[4] We can see this through a simple thought experiment. Imagine a complete model, estimated on a friendship network, that found positive and significant similarity effects on ISEI but not on ethnicity. This would suggest that actors did not consider ethnicity when making friendship choices. Now imagine that, in a nested model approach, we re-estimate the model without the ISEI similarity effect to test for the role of direct socioeconomic selection on aggregate levels of ISEI homophily. It may be the case that the ethnicity parameter turns positive and significant in this altered model due to the correlation of ISEI and ethnicity (meaning the confounding effect of ISEI similarity on ethnic similarity was accounted for by the full model). This altered model would represent actors as having a preference for ethnic similarity, and this preference would contribute to generating ISEI homophily due to the correlation between ISEI and ethnicity. Nevertheless, the full model, better informed about the correlation between the two attributes, would disagree that a mechanism of ethnic selection exists and, therefore, that it may generate socioeconomic homophily.

simulated outcomes are attributable to this process.[5] By contrast, in the second case, the two DGPs are different regarding several parameters, so we cannot pinpoint the difference in simulated outcomes to a single low-order process. However, the altered model will simulate networks whose target statistics are similar to the original network, except for the removed effect. Consequently, the differences in simulated outcomes are attributable to this statistic, i.e., a macro-level feature. This might be best thought of as looking at the structural constraints imposed by a given network feature on other features, with the simulated DGP remaining a "black box" that connects the imposed structural constraints (fixed statistics) and the macro-level outcome of interest.

For this reason, the two methods are not answering the exact same question. Say that we want to look at the impact of the transitivity effect on socioeconomic homophily. The first approach (parameter manipulation) looks at how the process of closing triplets generates homophily, whereas the second (nested models) focuses on how the number of closed triplets in a network constrains homophily. The first comes closer to the logic of an ABM, with a focus on the generative processes that underlie the network; whereas the second is more agnostic about the exact specification of these processes but is interested in robust relationships between macro features in distributions of networks, with a framework closer to that of regression models. I believe the two approaches are defensible, as SAOM integrates elements of both ABM and regression models (see Snijders and Steglich, 2015 for a discussion). The reason I chose not to re-estimate parameters here is that I wish to push the generative reasoning as far as possible, coherent with my theoretical interest for the low-order behavior of students (selecting similar friends, closing triplets, etc.) more so than for the corresponding aggregate features (levels of segregation of the network, of transitivity, etc.).

The second critical particularity of contribution scores is that the macro-level statistic they investigate corresponds to a target statistic of the estimated model. By contrast, previous works have considered macro features not directly included in the model specification, such as the average path length of a network or its degree distribution. The choice of either option depends on what the analyst is interested in explaining. In the present case, I wish to study the impact of direct socioeconomic selection (choosing friends similar to oneself) on aggregated levels of homophily (having friends similar to oneself): the symmetry in construction between process and outcome is intrinsic to this particular research question.[6] However, this has further implications on the issue of whether parameters should be re-estimated. Indeed, when the investigated feature is a target statistic of the model, re-estimating the corresponding parameter will always ensure that the predicted statistic is equal to the observed one, making the entire approach simply senseless. For this reason, not re-estimating parameters may be preferable when investigating a target statistic specifically.

3.3.5. Application to the present study

I consider nine potential inducers of socioeconomic homophily, that is, nine hypothesized relational processes represented by one or more model parameters.

Homophilic selection based on socioeconomic origin, grades, ethnicity, and gender is represented in the SAOM by similarity effects on the corresponding attributes. The first three should contribute to socioeconomic homophily: socioeconomic selection is the most direct way to induce socioeconomic homophily, whereas academic and ethnic selection should contribute to it through a compositional effect, as these attributes correlate to socioeconomic origin. Gender homophilic selection, on the other hand, should not contribute to socioeconomic homophily because gender and socioeconomic origin are unrelated.

Propinquity is represented by the similarity effects on students' place of residence and classroom. To test for the effect of residential propinquity, three parameters are jointly set to 0: the walking distance between students' homes, the squared distance, and the similarity effect of former primary school. Indeed, primary schools depend on students' place of residence and can be expected to impact friendship formation above and beyond the exact distance between their homes. As for the effect of classroom propinquity, it is tested by setting the two similarity parameters for the current and previous wave's classrooms to 0.

The contribution of special tracks is tested by setting to 0 the two parameters of special track similarity and regular track similarity (only in the schools Paris 1 and Savoie 1). Note that this likely captures a mixture of propinquity and homophilic selection: students in these tracks have more opportunities for contact among themselves (shared classes), but they are also well-identified by their peers, with an institutional label often acting as a stigma (see Palheta, 2012 for a discussion of special tracks in French middle schools).

The contribution of reciprocation (the process by which individuals reciprocate nominations) is tested by setting the reciprocity parameter to 0. Finally, the contribution of transitive closure (the process by which individuals close transitive triplets) is tested by jointly setting to 0 the two transitivity parameters (Forward-Forward and Forward-Back), as well as the interaction of the first (Forward-Forward) with reciprocity. Note that, by definition, this interaction effect pertains to both reciprocity and transitivity; however, the fact that it is a triadic effect with a 'gwesp' form, as well as the fact that it is functionally very close to a 3-cycle effect (Block, 2018), makes it conceptually closer to triadic closure than to reciprocation.

4. Results

4.1. Descriptive results: permutation tests

To test for the presence of socioeconomic homophily in the four schools, I compare the observed average absolute difference in ISEI scores among ties to that obtained in randomly permuted networks (500 permutations). Table 1 gives the average ISEI difference among ties in the observed and permuted networks. Fig. 1 shows the density curves of ISEI differences among ties, again in observed and permuted networks.

Table 1 shows that observed values are significantly lower than those in permutations for three schools out of four (all but Savoie 2). Moreover, for these three schools, we can see in Fig. 1 that observed densities tend to be higher than the permuted reference for ties with small ISEI differences (i.e., ties among socioeconomically similar students) and lower for ties with large differences. By contrast, in Savoie 2, observed values are virtually equal to the permuted reference regarding both

Table 1
Average ISEI Differences in Friendship Ties.

School	Observed	Mean in Permutations	P-Value
Paris 1	1.05	1.37	0.00
Paris 2	1.05	1.15	0.00
Savoie 1	0.92	1.02	0.00
Savoie 2	1.11	1.12	0.22

[5] Note that with contribution scores, the density parameter is re-estimated, which means that the DGP does change slightly beyond the parameter of interest. However, it makes theoretical sense to consider that the density parameter does not represent a real-life process (there is probably no such thing as a baseline propensity to make any sort of friends) but is meant to center other effects. Thus the simulated changes in the network structure should still be attributed to the mechanism represented by the omitted parameter.

[6] A roundabout way to answer this question while *not* resorting to a target statistic may be to measure socioeconomic homophily through a slightly different measure than the sum of ISEI scores (e.g., an index of segregation like Moran's coefficient). However, I do not believe that this fundamentally changes the issue. Either this alternative measure is very closely related to the target statistic, in which case both essentially hold the same information, and the results will be practically identical; or it is more distantly related to it, in which case they are in fact measuring different network features and the test will not answer the question intended by the analyst.

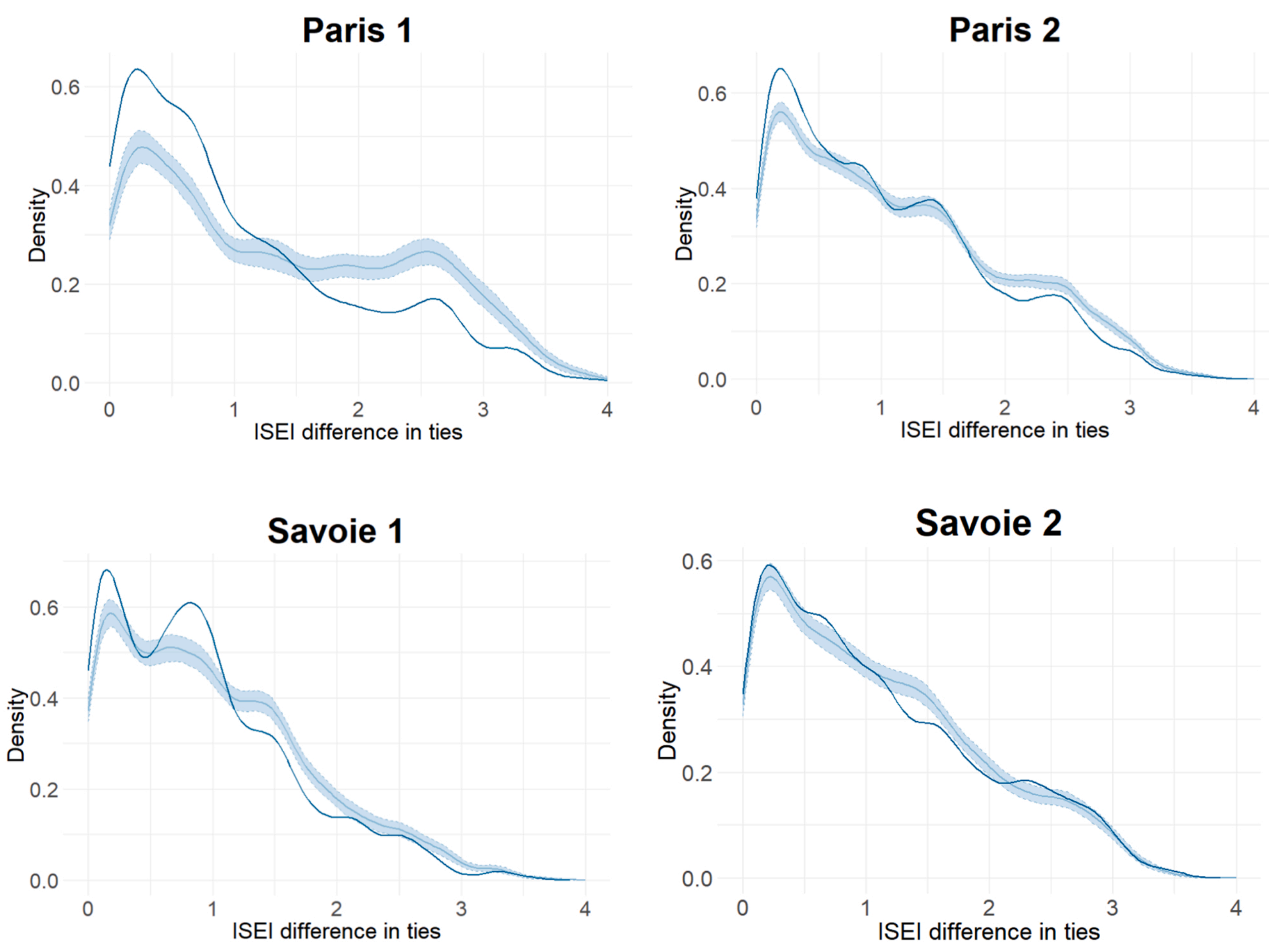

Fig. 1. Density Curves of ISEI Differences in Friendship Ties. Observed networks = solid dark blue line. Average in permutations = light blue line. 95% confidence interval in permutations = light blue shaded array.

mean values (Table 1) and the density distribution (Fig. 1).

To get a sense of the magnitude of socioeconomic homophily, it is possible to compute the log-odd change in the probability of observing a tie per additional point of the ISEI difference between nodes (which is equivalent to running a bivariate logistic regression at the tie level). Estimated log-odd coefficients are − 0.38 (Paris 1), − 0.17 (Paris 2), and − 0.21 (Savoie 1). The typical distance between a student whose parents are professionals (standardized ISEI approximately equal to 1.3) and a student whose parents are non-skilled workers (standardized ISEI approximately equal to −1) is approximately 2.3 points. Therefore, and converting the log-odd coefficients to odd ratios, we find that a tie between two students from a working-class and an upper-class background is approximately 2.4 (Paris 1), 1.4 (Paris 2), and 1.6 (Savoie 1) less likely to happen, rather than not, than a tie between same-background students.[7] From these figures, it is clear that homophily is much higher in Paris 1 compared to the other schools.

Explaining the difference in levels of homophily across the four schools is outside the scope of this article. However, there are reasons to believe that Savoie 2, the only school not exhibiting any homophilic pattern, is a relatively exceptional case, mainly because of a stronger selection bias in the student population (the school has a selective admission policy, such that the working-class households attending it are likely selected on unobserved properties that make them culturally more similar to middle- and upper-class households). Furthermore, there is a little bit of socioeconomic homophily in Savoie 2 as well, though very slight, which appears when considering stronger and rarer relationships (investigated through a different name generator in this same study; see Chabot (2021) for details, including ethnographic insights on each school). Consequently, if we attempt an out-of-sample prediction, we might expect the majority of diverse schools to exhibit patterns closer to those of Paris 1, Paris 2, and Savoie 1, that is, to exhibit socioeconomic homophily.

4.2. SAOM estimation

We now turn toward assessing the contribution of different tie-formation processes to the emergence of socioeconomic homophily. I focus here on the three schools where some of this homophily was observed, thus leaving Savoie 2 aside. Table 2 gives the results of the SAOM estimation in these three schools. Only the effects whose contribution scores I will compute are displayed here. The complete models with all included controls are in Appendix 6, together with a similar SAOM estimated for Savoie 2 and goodness-of-fit assessments and time-heterogeneity tests for all schools. Additionally, descriptives about the correlation of the tested covariates to socioeconomic origin are in Appendix 5.

Most of the effects presented here are statistically significant, suggesting that there are homophilic selection or propinquity effects for most of the considered covariates. Nevertheless, some exceptions are worthy of attention.

The coefficient of the ISEI similarity effect is positive and significant in Paris 1 and Paris 2. Given that the effect sizes are relatively large; that

[7] Paris 1: exp(0.38 *2.3) ∼= 2.4. Paris 2: exp(0.16 *2.3) ∼= 1.4. Savoie 1: exp(0.21 *2.3) ∼= 1.6.

Table 2
Results for the SAOM Estimation.

Effect Name	Paris 1 est.	sgn.	s.e.	Paris 2 est.	sgn.	s.e.	Savoie 1 est.	sgn.	s.e.
outdegree (density)	-2.88	***	(0.14)	-2.71	***	(0.05)	-2.84	***	(0.07)
reciprocity	2.41	***	(0.19)	2.74	***	(0.07)	2.42	***	(0.12)
gwespFF (transitivity)	0.89	***	(0.12)	0.46	***	(0.06)	0.66	***	(0.11)
gwespFB (transitivity)	0.33	***	(0.13)	0.65	***	(0.07)	0.56	***	(0.14)
gwespFF * reciprocity	-0.59	***	(0.10)	-0.58	***	(0.03)	-0.67	***	(0.09)
isei similarity	0.21	*	(0.12)	0.14	***	(0.04)	-0.04		(0.08)
same ethnicity	0.15	**	(0.06)	0.17	***	(0.02)	0.18	***	(0.05)
grades similarity	0.53	***	(0.17)	0.22	***	(0.06)	0.20	*	(0.11)
same gender	0.46	***	(0.06)	0.30	***	(0.02)	0.28	***	(0.03)
same classroom – current wave	0.48	***	(0.07)	0.81	***	(0.02)	0.71	***	(0.04)
same classroom – previous wave	0.17	**	(0.06)	0.10	***	(0.02)	0.19	***	(0.03)
special track – same * special track	0.17		(0.17)	-	-	-	-0.15		(0.16)
special track – same * normal track	0.56	***	(0.13)	-	-	-	0.41	***	(0.12)
same primary school	0.04		(0.06)	0.00		(0.02)	0.26	***	(0.04)
residential distance	-0.29		(0.26)	-0.28	***	(0.05)	-0.06	**	(0.02)
residential distance – squared	0.29		(0.19)	0.09	***	(0.02)	0.005	**	(0.002)

Note: “est.” = estimated coefficient; “sgn.” = statistical significance; “s.e.” = standard error.
For continuous attributes (isei and grades), the SIENA algorithm automatically centers the absolute difference among nodes, rescales it on the (0,1) interval and reverses it (i.e. substract it to 1). As a result, homophily is indicated by positive coefficients. However, for residential distance, the covariate, though continuous, has not been transformed, so residential homophily is indicated by negative coefficients (when the distance between students increases, the chances of a tie decrease).
Additional effects not shown here include: for structural effects, reciprocity, 3-cycles (Paris 2 and Savoie 1), reciprocity*transitivity, and two-paths (Paris 2); for degree-related effects, indegree popularity, outdegree activity and outdegree popularity (linear or squared version depending on the school); and for covariate-related effects, emission and reception terms for isei, ethnicity, grades and gender, plus a single emission effect for special tracks (Paris 1 and Savoie 1), and a time-interaction dummy for the classroom of the previous year. The decay for the gwesp parameters depends on the schools. See Appendix 6.
* $p < 0.1$;
** $p < 0.05$;
*** $p < 0.01$

the models include many relevant covariates correlated to socioeconomic origin (ethnicity, grades, place of residence, tracks, and classrooms); and that structural effects offer a certain amount of protection against omitted variables; it seems relatively unlikely that these effects would only be the result of unobserved confounders. Thus there seems to be homophilic selection based on socioeconomic origin in the two Parisian schools. In Savoie 1, however, the coefficient is negative, close to 0, and statistically insignificant. Therefore, there does not seem to be any socioeconomic selection in Savoie 1, which implies that other relational processes can entirely account for the socioeconomic homophily of the friendship networks.

In Paris 1, residential distance does not seem to play a role in friendship formation, either directly or through the former primary school. In Paris 2, there is an effect of residential distance, but not of primary schools. As for Savoie 1, all three parameters are statistically significant. The absence of an effect of residential distance in Paris 1 is likely due to its small size and location in a dense neighborhood: Most students live at a walking distance from one another, and even when they do not public transportation is abundant (subway and bus). By contrast, Paris 2 covers a broader residential area, and public transportation, though still available, is less dense. As for Savoie 1, it is located in a semi-rural area, with little public transportation and large distances between students' homes, so we might expect residential propinquity to matter even more there.

Finally, it is interesting to note that, in Paris 1 and Savoie 1, the similarity effects for special tracks are statistically significant for students in the regular track (i.e., the large majority of the population) but not for those in the special tracks. Students from the regular tracks appear reluctant to nominate those in the special tracks, but not the other way around. Therefore, there might be a status asymmetry, with a stigma associated with students from the special tracks. Interestingly, this pattern is similar in the elitist track of Paris 1 and the pre-vocational one of Savoie 1.

4.3. Contribution scores

The SAOM coefficients of Table 2 indicate which relational processes are active in the evolution of the networks over time. However, they do not measure the importance of these processes regarding socioeconomic homophily precisely. For this, I resort to the contribution scores presented in Section 3.3. Remember that I compute these scores by setting one or several coefficients to 0, then simulating networks from this altered model. I only test the parameters significantly different from 0 in the SAOM estimation. Indeed, by definition, any null effect will have a null contribution score. Therefore, if we cannot be confident that a “true” effect size is different from 0 (following the usual definition of statistical significance), then we cannot be confident either that the corresponding “true” contribution is different from 0.

Table 3 gives the contribution scores under the different simulation scenarios. These scores give the proportion change (or percentage change when multiplied by 100) in socioeconomic homophily predicted by the model when setting to 0 the considered parameters. Furthermore, Fig. 2 shows the distributions of the ISEI similarity statistic in the 1000 networks simulated for each scenario.

The parameters exhibiting consistently high contribution scores across the three schools are reciprocity and transitivity. When the reciprocity parameter is set to 0, the average ISEI homophily observed in the simulations drops by 43% (Paris 1), 52% (Paris 2), and up to 61% (Savoie 1). As for transitivity, setting the three corresponding parameters to 0 results in 31% (Paris 1), 60% (Paris 2), and 67% (Savoie 1) less ISEI homophily.

A word of caution here: Remember that the scores refer to a total homophily that is not the same in the three schools. There is more homophily in Paris 1 than in the other schools (Section 4.1), so the fact that transitivity generates a proportionally smaller part of homophily in this school does not mean that it generates less homophily in absolute

Table 3
Contribution scores.

		Predicted change in ISEI homophily		
Scenario Name	**Parameters set to 0**	Paris 1	Paris 2	Savoie 1
recip	reciprocity	-0.43	-0.52	-0.61
gwesp	transitivity (FF, FB and FB*recip)	-0.31	-0.60	-0.67
isei	isei similarity	-0.32	-0.47	n. sign.
ethnicity	same ethnicity	-0.03	-0.14	-0.22
grades	grades similarity	-0.05	0.00	0.00
gender	same gender	0.07	-0.04	-0.04
classroom	current year classroom + past year classroom	-0.11	0.17	-0.03
special track	same * special track + same * normal track	-0.14	-	0.05
primary_walk	same primary school + residential distance + residential distance squared	n. sign.	0.02	-0.19

Note: "n. sign." indicates that the corresponding parameter(s) in the SAOM are not significantly different from 0 (see Table 2). Therefore, the corresponding contribution scores are assumed to be either null or close from null.
Example of interpretation: in Paris 1, when the gwesp parameter is set to 0, simulated networks exhibit on average 46% less socioeconomic homophily than the observed network.

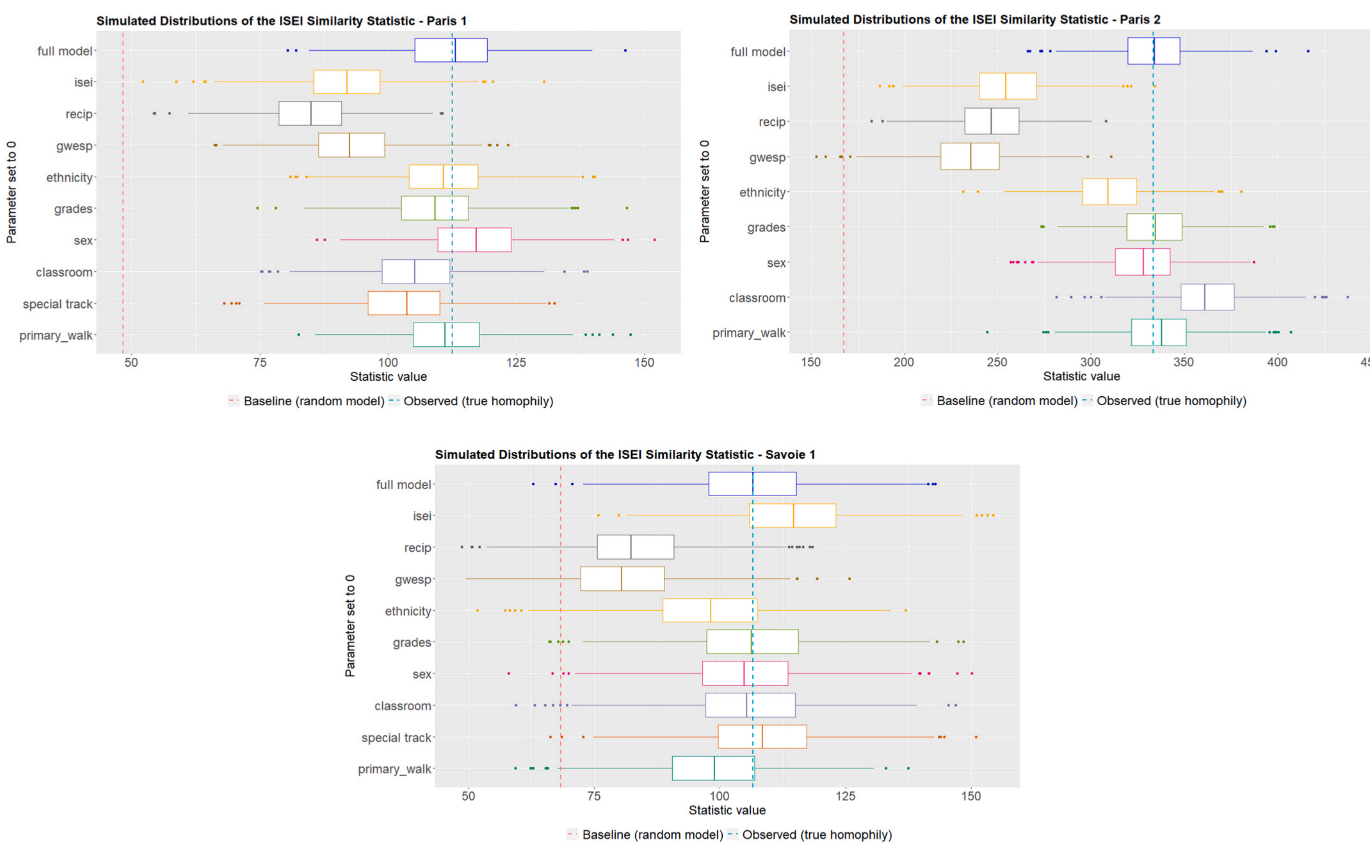

Fig. 2. Distributions of Simulated Statistics for ISEI Homophily under the Different Scenarios. Note: boxes represent the values of the ISEI statistic in the simulated networks. The blue vertical line corresponds to the true observed value of the statistic, and the red line to the average value observed under the null distribution. Therefore, the total amount of socioeconomic homophily observed empirically is represented by the distance between the red and blue lines, and the reduction in the ISEI similarity statistic is the distance between the blue line and the center of the box. Contribution scores are equal to the difference between the mean simulated value of the statistic and its observed value, divided by the difference between the observed and null values (see Eq. 4). On the Figure, this corresponds to the distance between the center of the box (average simulated value) and the blue line (observed value), divided by the distance between the blue and red line (null value).

terms.[8]

In Paris 1 and 2, the contribution of ISEI similarity is − 0.32 and − 0.47, respectively. Therefore, assuming that there is no direct homophilic selection whatsoever, the model predicts that socioeconomic homophily will be divided by 2 or 3; or, conversely, that between half and two-thirds of this homophily can be explained by other relational processes. As for Savoie 1, we already saw from the SAOM of Table 2 that the ISEI similarity parameter is non-significant, meaning that other effects can entirely account for socioeconomic homophily.

Savoie 1 is also the only school where residential propinquity notably impacts socioeconomic homophily. The joint contribution score of residential distance and the former primary school is − 0.19, meaning that homophily drops by 19% when these effects are set to 0. As

[8] One could look at a substantial measure of homophily in the simulated networks (e.g., through odd ratios). However, comparing substantial effect sizes across networks with different densities and structures is not straightforward, so I do not engage with this here (given that my prime focus is not on comparing levels of homophily across the schools).

mentioned in Section 4.2., this probably comes from the rural location of the school, which implies that there is less public transportation and larger distances among students' homes. Supplementary analyses (not shown) indicate that this contribution is explained almost entirely by the effect of the former primary school, whose impact on socioeconomic homophily is much more decisive than residential distance itself.

Other covariates have relatively low scores overall. The 'classroom' scenario exhibits a score of − 0.11 in Paris. In Paris 2 and Savoie 1, classrooms have virtually no inducing effect (scores of 0.17 and −0.03, respectively; see below for the interpretation of positive scores). Looking at descriptives (Appendix 5), this is easy to explain: Paris 1 is the school where differences across classrooms in students' average ISEI are most pronounced.

The special track of Paris 1 also has an inducing effect on socioeconomic homophily (score of −0.14). This contribution is relatively modest, given that this track welcomes students from very high socioeconomic origins compared to the rest of the school. Most likely, they are too few to greatly impact the socioeconomic homophily of the entire network (~15 students out of 90). As for Savoie 1, the score of special track similarity is positive and very small, at 0.05.

Ethnic similarity exhibits a noticeable score in Paris 2 (−0.14) and even more so in Savoie 1 (−0.22), but a small one in Paris 1 (−0.03). This is surprising because the individual-level correlation between ethnicity and socioeconomic background is stronger in Paris 1 than in the other schools (Appendix 5), and the SAOM estimations found a significant effect of ethnic similarity in all three schools (Table 2). It may be due to a larger effect of ethnic similarity in Paris 2 and Savoie 1 than in Paris 1 (which is difficult to say for sure based on the SAOMs, since the size of coefficients is not directly comparable across models), but also to low-order peculiarities of the distribution of ethnic backgrounds in the networks, which are not immediately apparent from descriptives but that are impactful during the simulations nevertheless (e.g., same-ethnic ties might be more clustered in certain schools, thus more sensitive to aggravation processes linked to transitive closure).

Grades similarity appears to be a weak inducer, with a score of merely − 0.05 in Paris 1 and 0.00 in Paris 2 and Savoie 1. As for gender similarity, it exhibits a score of − 0.04 in Paris 2 and Savoie 1 due to a slight socioeconomic imbalance between boys and girls (girls are from slightly higher origins on average). In Paris 1, on the other hand, the score is positive at 0.07.

Finally, a word on positive scores (gender in Paris 1 and classroom in Paris 2). They indicate that setting a parameter to 0 increases the predicted amount of socioeconomic homophily. By extension, the considered parameters induce socioeconomic *heterophily* under the full model. However, this does not necessarily mean that the change statistics of the two effects are negatively correlated; for example, the positive score for gender in Paris 1 (0.07) does not imply that same-gender dyads have a lower socioeconomic similarity on average (in fact, the dyadic correlation is positive, though almost null, at 0.02). Instead, it is a formal implication of SAOM that setting one effect to 0 will slightly affect the impact of all other modeled effects (since all effects effectively interact throughout the simulation procedure). In the case of positive scores, a substantial interpretation might be that tie-formation processes tend to constrain one another: the more processes of tie-formation there are, the less "space" each one has to be expressed (figuratively speaking). For example, students being indifferent to their friends' gender would result in them having more latitude to select friends based on their socioeconomic origin.

5. Discussion

5.1. Inducers of socioeconomic homophily: direct selection, minor inducers, and endogenous aggravation

In this article, I examined friendship networks in four French middle schools with high levels of socioeconomic diversity. Three schools out of four exhibit socioeconomic homophily, in the sense that ties are more likely to be observed among students from similar socioeconomic origins. The magnitude of this homophily, however, markedly varies across schools, as it appears to be strong in Paris 1 but relatively moderate in Paris 2 and Savoie 1.

How important is socioeconomic selection – the relational process by which individuals select friends similar to them in socioeconomic terms – in explaining this observed homophily? In one out of the three schools (Savoie 1), the estimated Stochastic Actor-Oriented Model showed no sign of students selecting their friends based on socioeconomic origin. As for the other two schools, simulations derived from the estimated SAOMs generate about a third (Paris 1) or half (Paris 2) less homophily under the assumption that there is no selection. Additionally, one should note that these are likely inflated estimates: by construction, the parameter of socioeconomic selection is the one that will capture most of the impact on socioeconomic homophily of omitted covariates (as the model might not account for all processes of interest). If we were to control for additional processes, this might explain away yet another part of the effect. Therefore, assuming that the estimated SAOM is a reasonable representation of the true data-generating process, homophilic selection appears as an important process driving homophilic tie formation, but not in all schools, as well as it is far from being the only relevant explanation.

The other major relational processes contributing to the emergence of socioeconomic homophily appear to be reciprocation and transitive closure. In all three schools, scenarios with the reciprocity parameter set to 0 exhibit much lower socioeconomic homophily, from 43% to 61% less. Similarly, setting the transitivity parameters to 0 results in homophily dropping by 31–67%. These figures are coherent with previous studies that found transitive closure to reinforce initial levels of homophilic selection, both in empirical contexts (Goodreau et al., 2009) and through stylized simulations (Asikainen et al., 2020; Foster et al., 2011). Crucially, neither reciprocation nor transitive closure can, on their own, induce homophily: they only reinforce an initial imbalance in tie probabilities. In other words, their contribution to socioeconomic homophily is due to the aggravation of the impact of other relational processes.

Conceptually, these other processes are of two kinds. First, reciprocation and transitive closure can aggravate the impact of the processes represented by other model parameters (say, ethnic homophilic selection or classroom propinquity). The studies I just mentioned focus on such effects since they use cross-sectional data and simulations.

Second, in the case of a model of network evolution like SAOM, reciprocation and transitive closure tend to increase the inertia of the homophily found in past observations of the network. Indeed, to the extent that homophilic ties are more reciprocal or clustered than non-homophilic ones in the observed network at time t, the reciprocity and transitivity parameters in SAOM will allocate ties more frequently in the homophilic regions of the network when simulating fictive $t + 1$ networks. Part of the homophily in t carries over to $t + 1$ through this mechanism. In terms of interpretation, this means that reciprocation and transitive closure can induce a ratchet effect: once homophilic friendship groups form, they become unlikely to dissolve. In that case, the aggravated relational processes are those that gave birth to the initial network of wave 1, some of which may be similar to the processes captured by the SAOM estimation, and some of which may pertain to unobserved factors. In that regard, it is good to remember that SAOM, and thus contribution scores, are only informative about the processes that drive the network evolution from wave 1 onward. One may reasonably assume that the processes underlying the initial network are similar to these, but this needs not be the case.

Most other parameters exhibit relatively modest contribution scores, at least when considered in isolation: classrooms and tracks in Paris 1, ethnicity in Paris 2, or ethnicity and former primary schools in Savoie 1 only contribute to about 10–20% of the total socioeconomic homophily each. Nevertheless, in scenarios with no socioeconomic selection, the models still predict the emergence of a large share of socioeconomic homophily. Thus these weak contributors seem sufficient to "tilt" the network structure toward socioeconomic homophily, with reciprocation and transitive closure then aggravating this initial imbalance. Therefore, even in the absence of direct homophilic selection, it appears that homophily can emerge as the joint product of several minor inducing processes combined and reinforced by endogenous network processes.

These results echo previous findings that strong segregation can emerge from weak individual inclinations toward similarity. This is notably the case in Schelling's famous model of spatial segregation (Schelling, 1969), but the same idea has been extended to homophily in social network contexts (Stadtfeld, 2018). What we see here (particularly from Savoie 1) is that it is not even necessary for this weak selection effect to pertain to the attribute of interest directly; homophilic selection on correlated attributes, as well as propinquity processes, can be enough for a noticeable homophily to emerge through compositional effects. With that said, a point of caution has been raised by Block (2018), who argued that there might often be negative interaction effects between homophilic selection and endogenous network processes and that this might attenuate the impact of the latter on the aggregate homophily. Future works could use contribution scores to investigate such interactions and further decompose the processes through which homophily emerge.

5.2. Contribution scores and model analysis for SAOM

Using simulation models to test fictional scenarios and pinpoint the role of specific relational processes in the emergence of network structure is a complex task, and the literature has yet to reach a consensus on how exactly this should be done. I tried to contribute to this effort by proposing a particular take on this question. *Contribution scores* have two noteworthy properties compared to previous approaches: they fix the value of the full model's parameters rather than re-estimating them, coherent with a generative framework where higher-order structures are contingent outcomes of low-order processes; and they investigate a target statistic of the model. I believe that this was the most appropriate approach given my data and research questions, but this may differ across studies. The fact that the results obtained here resonate with existing theories in Social Network Analysis regarding the aggravation of homophilic selection through reciprocation and transitive closure is encouraging as per the validity of the approach. However, further research is needed to assess its scope of validity and potential biases. One may in particular compare the results and implications of different methods applied to the same data set.

This also echoes ongoing discussions about the way to measure effect size in SAOM. Indeed, parameters' values are hard to interpret and depend on network size as well as model specification, so researchers typically struggle to know when an effect should be considered "large" or "small". Indlekofer and Brandes (2013) measure of "relative importance" approach this issue by attempting to single out the impact of a parameter at a given mini-step in the SAOM simulation. Contribution scores follow a different logic: They envision the impact of parameters at the level of network aggregate features. At the moment, the scores are not truly a measure of relative importance nor effect size, because they only consider a single aggregate feature at once – perhaps this might be described as a "targeted effect size" measure. However, a possible direction for research might be to move from targeted to generic effect size while still following the same logic, by considering the impact of an estimated parameter on a set of multiple network features.

Another point of particular interest pertains to positive scores. The parameters whose change statistic is not correlated whatsoever to that of the feature of interest tend to exhibit positive contributions, that is, a heterophilic effect. This was for example the case of gender similarity in certain schools.[9] At a formal level, this is an implication of how SAOM works. However, in substantial terms, it may be taken as appropriately modeling the fact that any relational process tends to "take some space" in actors' decisions, thus attenuating the impact of other processes. Thus, positive scores may capture the decreasing marginal relevance of multidimensional similarity, as theorized by Blau (1994) or Block and Grund (2014). It is an open discussion whether interpreting positive scores through this perspective is valid and appropriate. More broadly, any SAOM contains what may be called "implicit interactions" between all model parameters, because of the underlying multinomial model as well as the simulation procedure. Contribution scores can help detect these implicit interactions, though more research is needed to understand to which extent they capture real-world processes or are merely model artifacts.

Finally, a word on the fact that the scores do not sum up to -1 over all parameters. As mentioned in Section 3.3, this appropriately reflects the fact that there is some overlap between the effect of different parameters (both due to the dependency of network observations and the logit link function in SAOM's multinomial choice model). However, it also highlights a deeper conceptual issue: the impact of a given process on the network structure is always dependent on other processes. In that sense, even when the parameters of the estimated model are not re-estimated in the altered scenario, part of the contribution of the tested parameter nevertheless pertains to how its absence modifies the expression of other parameters. Consequently, the contribution of a given parameter – and thus of the substantial process it represents – should always be interpreted conditional on the rest of the DGP, i.e., in the context of a particular network and model. To better reflect this fundamental property of network processes, it might be interesting to divide contribution scores by the sum of the scores of all model parameters, similar to what Indlekofer and Brandes (2013) proposed for their measure of relative importance of effects in SAOM. This would result in "relative contributions" where the contribution of each parameter is only compared to that of other parameters, instead of being expressed in a substantial metric of homophily (or any other statistic). Arguably, this would further highlight the fact that contributions are always conditioned on a given DGP. It would also be particularly interesting to compare relative and absolute scores. Intuitively, their difference (i.e., how far away from -1 the sum of all scores is) should reflect the extent to which the low-order mechanisms enacted by the model interact and synergize. However, in a model with many parameters, this might come at a high computational cost.

6. Conclusion

The results presented here suggest that there generally is socioeconomic homophily among middle-school students that attend socioeconomically diverse schools. This homophily, it seems, does not primarily result from a propensity to select same-background friends. Though this propensity may exist in certain schools, a significant share of homophily is also explained by compositional effects linked to ethnicity, place of

[9] Note that the nested model approach (i.e., re-estimating parameters) can also result in such positive scores. I tested this for the gender similarity parameter in Paris 1 by comparing an estimated model without ISEI similarity to a second model without both ISEI similarity *and* gender similarity (so that the two models would allow the simulated networks to have different values of the ISEI statistic). The predicted ISEI homophily was indeed larger for this second specification, by 5% (to be compared with the 7% found in Table 3).

residence, or curricular options. Most importantly, endogenous network processes – reciprocation and transitive closure – strongly aggravate the impact of other relational processes, resulting in aggregated levels of socioeconomic homophily that far exceed individuals' homophilic inclinations.

Further research is needed to understand better the role of different factors in the emergence of socioeconomic homophily among school students. In that regard, three directions (at least) seem worthy of exploring. First, not all relevant inducing processes have been considered in this article: for example, I could not examine the impact of out-of-school activities, which act as foci where students can meet certain friends, nor that of parents, which exert a certain amount of control over the sociability of their children. Second, the finding that there seems to be homophilic selection based on socioeconomic origin (in two out of four schools) does not tell us how this process operates precisely. Are students aware of their peers' origins? Do they rely on external signals of socioeconomic status, such as clothing? Or is socioeconomic selection mediated by similarity in psychological traits, such as interests, tastes, personality, or values? Finally, properties of the school context likely impact the strength of homophilic selection, making students' inclination toward similar others more or less salient. These are known as "contextual moderators" or "ecological moderators". For example, hostile environments where students feel insecure may reinforce their tendency to bond with similar peers as they seek a sense of security and predictability in their relationships (McFarland et al., 2014).

Although the results presented here cannot directly predict the effect of an intervention, they do support the idea that the studied adolescents did not have strong discriminatory attitudes toward different-background peers. There is a long tradition of works in sociology and social psychology arguing that racial homophily is driven by contextual incentives more than intrinsic preferences (e.g., Allport, 1954; Blau, 1994; Wimmer and Lewis, 2010); and the same might be true of socioeconomic homophily. In turn, this suggests that school administrations and public policies have some latitude to favor diversity in students' friendships by desegregating meeting contexts like schools, classrooms, and neighborhoods.

Declaration of Competing Interest

None.

Acknowledgments

Part of this research was conducted as a Ph.D. candidate at the European University Institute (EUI, Florence, Italy), and later on as a postdoctoral researcher at the French Institute for Demographic Studies (INED, Paris, France). The author thanks Tom A.B. Snijders, Christian Steglich and the participants to the SIENA Advanced Users Meetings from the Groningen Winter School, as well as Marion Hoffman and Daniel Gotthardt, for their feedbacks on earlier versions of the article.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.socnet.2023.09.002.

References

Allport, G.W., 1954. The Nature of Prejudice. Addison-Wesley Publishing Company,.
Asikainen, A., Iñiguez, G., Ureña-Carrión, J., Kaski, K., Kivelä, M., 2020. Cumulative effects of triadic closure and homophily in social networks. Sci. Adv. 6 (19), eaax7310.
Baerveldt, C., Van Duijn, M.A.J., Vermeij, L., Van Hemert, D.A., 2004. Ethnic boundaries and personal choice. Assessing the influence of individual inclinations to choose intra-ethnic relationships on pupils' networks. Soc. Netw. 26 (1), 55–74.
Ball, S.J., Vincent, C., Kemp, S., 2004. "Un agréable mélange d'enfants. ": prise en charge de la petite enfance, mixité sociale et classes moyennes. Educ. Et. sociétés no 14 (2), 13–31.
Blau, P.M., 1994. Structural Contexts of Opportunities. University of Chicago Press,.
Block, P., 2015. Reciprocity, transitivity, and the mysterious three-cycle. Soc. Netw. 40, 163–173.
Block, P., 2018. Network evolution and social situations. SocScience 5, 402–431.
Block, P., Grund, T., 2014. Multidimensional homophily in friendship networks*. Netw. Sci. 2 (2), 189–212.
Block, P., Koskinen, J., Hollway, J., Steglich, C., Stadtfeld, C., 2018. Change we can believe in: Comparing longitudinal network models on consistency, interpretability and predictive power. Soc. Netw. 52, 180–191.
Bourdieu, P., 2016. La Distinction: Critique sociale du jugement. Minuit,.
Chabot, T., 2021. From diversity to mixing? Socioeconomic homophily in French desegregated middle schools (Ph.D. thesis). European University Institute. https://cadmus.eui.eu/handle/1814/73517.
Coleman, J.S., 1961. The Adolescent Society: The Social Life of the Teenager and Its Impact on Education. Free Press of Glencoe,.
Coleman, J.S., 1995. Equality of Educational Opportunity (Coleman) Study (EEOS), 1966. Inter-Univ. Consort. Political Soc. Res.
Coleman, J.S., Hoffer, T., 1987. Public and Private High Schools: The Impact of Communities. Basic Books,.
Duru-Bellat, M., Mingat, A., 1997. La constitution de classes de niveau dans les collèges; les effets pervers d'une pratique à visée égalisatrice. Rev. Fr. De. Sociol. 38 (4), 759.
Eder, D., 1985. The cycle of popularity: interpersonal relations among female adolescents. Sociol. Educ. 58 (3), 154.
Feld, S.L., 1981. The focused organization of social ties. Am. J. Sociol. 86 (5), 1015–1035.
Flashman, J., 2012. Academic achievement and its impact on friend dynamics. Sociol. Educ. 85 (1), 61–80.
Foster, D.V., Foster, J.G., Grassberger, P., Paczuski, M., 2011. Clustering drives assortativity and community structure in ensembles of networks. Phys. Rev. E. 84 (6), 066117.
Fujimoto, K., Snijders, T.A.B., Valente, T.W., 2018. Multivariate dynamics of one-mode and two-mode networks: Explaining similarity in sports participation among friends. Netw. Sci. 6 (3), 370–395.
Ganzeboom, H.B.G., De Graaf, P.M., Treiman, D.J., 1992. A standard international socio-economic index of occupational status. Soc. Sci. Res. 21 (1), 1–56.
Goodreau, S.M., Kitts, J.A., 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.
Handcock M.S., Hunter D.R., Butts C.T., Goodreau S.M., Krivitsky P.N., et al., 2016, statnet: Software Tools for the Statistical Analysis of Network Data.
Hunter, D.R., 2007. Curved exponential family models for social networks. Soc. Netw. 29 (2), 216–230.
Ichou, M., 2018. Les enfants d'immigrés à l'école: Inégalités scolaires du primaire à l'enseignement supérieur. Presses Universitaires de France,, Paris.
Indlekofer, N., Brandes, U., 2013. Relative importance of effects in stochastic actor-oriented models. Netw. Sci. 1 (3), 278–304.
Kruse, H., Smith, S., van Tubergen, F., Maas, I., 2016. From neighbors to school friends? How adolescents' place of residence relates to same-ethnic school friendships. Soc. Netw. 44, 130–142.
Lakon, C.M., Hipp, J.R., Wang, C., Butts, C.T., Jose, R., 2015. Simulating dynamic network models and adolescent smoking: the impact of varying peer influence and peer selection. Am. J. Public Health 105 (12), 2438–2448.
Lospinoso, J.A., Schweinberger, M., Snijders, T.A.B., Ripley, R.M., 2011. Assessing and accounting for time heterogeneity in stochastic actor oriented models. Adv. Data Anal. Cl. 5 (2), 147–176.
Malacarne, T., 2017. Rich friends, poor friends: inter–socioeconomic status friendships in secondary school. Socius 3, 2378023117736994.
Manzo, G., 2022. Agent-Based Models and Causal Inference. John Wiley & Sons,.
McFarland, D.A., Moody, J., Diehl, D., Smith, J.A., Thomas, R.J., 2014. Network ecology and adolescent social structure. Am. Sociol. Rev., 0003122414554001
McPherson, M., Smith-Lovin, L., Cook, J.M., 2001. Birds of a feather: homophily in social networks. Annu. Rev. Sociol. 27 (1), 415–444.
Mood, C., 2010. Logistic regression: why we cannot do what we think we can do, and what we can do about It. Eur. Socio Rev. 26 (1), 67–82.
Moody, J., 2001. Race, school integration, and friendship segregation in America. Am. J. Sociol. 107 (3), 679–716.
Oberti, M., Préteceille, E., Rivière, C., 2012, Les effets de l'assouplissement de la carte scolaire dans la banlieue parisienne. HALDE - défenseur des droits et DEPP - ministère de l'éducation nationale.
Octobre, S., Détrez, C., Mercklé, P., Berthomier, N., 2010. L'enfance des loisirs. Trajectoires communes et parcours individuels de la fin de l'enfance à la grande adolescence. Ministère de la Culture - DEPS,.
Oles, A., 2021, Query openrouteservice from R. Github. ⟨https://giscience.github.io⟩.
Palheta, U., 2012. Le collège divise. Appartenance de classe, trajectoires scolaires et enseignement professionnel. Sociologie 2 (4), 363–386.
R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing,, Vienna, Austria. ⟨www.r-project.org⟩.
Ripley, R., Snijders, T.A.B., 2020, Manual for SIENA version 4.0. 65.
Schelling, T.C., 1969. Models of segregation. Am. Econ. Rev. 59 (2), 488–493.
Shrum, W., Cheek, N.H., Hunter, S.MacD., 1988. Friendship in school: gender and racial homophily. Sociol. Educ. 61 (4), 227.
Smirnov, I., Thurner, S., 2017. Formation of homophily in academic performance: Students change their friends rather than performance. PLOS ONE 12 (8), e0183473.
Smith, S., Maas, I., van Tubergen, F., 2014. Ethnic ingroup friendships in schools: Testing the by-product hypothesis in England, Germany, the Netherlands and Sweden. Soc. Netw. 39, 33–45.

Snijders, T.A.B., 2001. The statistical evaluation of social network dynamics. Sociol. Methodol. 31, 361–395.
Snijders, T.A.B., 2017. Stochastic actor-oriented models for network dynamics. Annu. Rev. Stat. Its Appl. 4 (1), 343–363.
Snijders, T.A.B., Steglich, C.E.G., 2015. Representing micro–macro linkages by actor-based dynamic network models. Sociol. Methods Res. 44 (2), 222–271.
Stadtfeld, C., 2018. The micro–macro link in social networks. Emerg. Trends Soc. Behav. Sci. 1–15.
Stadtfeld, C., Amati, V., 2021. Network mechanisms and network models. In: Manzo, G. (Ed.), In *Research Handbook on Analytical Sociology*, pp. 432–452.
Steglich, C., Snijders, T.A., Pearson, M., 2010. Dynamic networks and behavior: Separating selection from influence. Sociol. Methodol. 40 (1), 329–393.
Steglich, C.E.G., Snijders, T.A.B., 2022. Stochastic Network Modeling as Generative Social Science. Edward Elgar Publishing,.
Wimmer, A., Lewis, K., 2010. Beyond and Below Racial Homophily: ERG Models of a Friendship Network Documented on Facebook. Am. J. Sociol. 116 (2), 583–642.
van Zanten, A.V., 2015. Choisir son école: Stratégies familiales et médiations locales. Presses Universitaires de France,.