The Emergence and Stability of Groups in Social Networks

Christoph Stadtfeldb,c*, Károly Takácsb,c, András Vörösa,d

a Chair of Social Networks, ETH Zürich, Switzerland
b The Institute for Analytical Sociology, Linköping University, Sweden
c Centre for Social Sciences, Computational Social Science – Research Centre for Educational and Network Studies (TK CSS-RECENS), Budapest, Hungary
d Department of Social Statistics and Mitchell Centre for Social Network Analysis, University of Manchester, United Kingdom

ARTICLE INFO

Keywords:
Definition of a group
Positive and negative ties
Structural balance
Stochastic actor-oriented models (SAOMs)
Agent-based models

ABSTRACT

An important puzzle in social network research is to explain how macro-level structures emerge from micro-level network processes. Explaining the emergence and stability of structural groups in social networks is particularly difficult for two reasons. First, because groups are characterized both by high connectedness within (group cohesion) and lack of connectedness between them (group boundaries). Second, because a large number of theoretical micro-level network processes contribute to their emergence. We argue that traditional social network theories that are concerned with the evolution of positive relations (forces of attraction) are not sufficient to explain the emergence of groups because they lack mechanisms explaining the emergence of group boundaries. Models that additionally account for the evolution of negative ties (forces of repulsion) may be better suited to explain the emergence and stability of groups. We build a theoretical model and illustrate its usefulness by fitting stochastic actor-oriented models (SAOMs) to empirical data of co-evolving networks of friendship and dislike among 479 secondary-school students. The SAOMs include a number of newly developed effects expressing the co-evolution between positive and negative ties. We then simulate networks from the estimated models to explore the micro-macro link. We find that a model that considers forces of attraction and repulsion simultaneously is better at explaining groups in social networks. In the long run, however, the empirically informed simulations generate networks that are too stylized to be realistic, raising further questions about model degeneracy and time heterogeneity of group processes.

1. Introduction

Important parts of human life are organized in social groups. In large social contexts, like schools, universities, workplaces, neighborhoods, and digital online communities, people form smaller groups in which they experience a higher level of communication, safety, belonging, interdependence, social norms, and psychological well-being as compared to the rest of the social context (Homans, 1950; Kadushin, 2002; Baumeister and Leary, 1995; Cartwright and Ed Zander, 1960; Turner et al., 1987; Berkman et al., 2000). How groups emerge and remain stable in various contexts is an important question in the social sciences, as group cohesion and group boundaries are linked to crucial societal outcomes such as segregation, group inequalities, differences in information access, and political polarization. The formation and stability of groups is further related to a central puzzle in sociology and social network research of how micro-level network processes aggregate into macro-level network structures (e.g., Stadtfeld, 2018).

Within social groups, individuals typically form strong positive ties, such as friendship (cf. Moody and White, 2003). The interpersonal relations that connect individuals who are members of different groups, however, tend to be weaker, neutral, or even negative (e.g., Granovetter, 1973; Tajfel and Turner, 1979; Stark et al., 2013). From a social network perspective, small groups thereby appear as network clusters—subsets of nodes in a positive social network that have a high density within (group cohesion) and are sparsely connected between (group boundaries). To distinguish groups based on this structural definition from broader concepts, we refer to them as structural groups in the following. Similar definitions have been linked to social groups in prior work (e.g., Reitz, 1988; Freeman et al., 1989; Freeman, 1992; Zeggelink et al., 1996).

* Corresponding author at: ETH Zürich, Chair of Social Networks, Department of Humanities, Social and Political Sciences, Weinbergstrasse 109, 8092 Zürich, Switzerland.
E-mail address: c.stadtfeld@ethz.ch (C. Stadtfeld).

The definition of structural groups that we use in this paper is a simplified representation of real social groups and, for example, assumes that individuals are members of just one group.

https://doi.org/10.1016/j.socnet.2019.10.008
Received 12 January 2018; Accepted 24 October 2019
Available online 18 November 2019
0378-8733/ © 2019 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/BY/4.0/).
The emergence of structural groups is difficult to link to micro-level social mechanisms in comparison to broadly explored macro-level outcomes such as, for example, specific types of degree distributions (Price, 1976; Barabási and Albert, 1999), or small-world features (Watts and Strogatz, 1998). First, because structural groups are relatively complex outcomes that are simultaneously characterized by group cohesion and group boundaries. Second, because multiple micro-level network processes contribute to their emergence and stability.

This article proposes a theoretical model in which multiple social network processes jointly explain the formation and stability of social groups, and illustrates its value through an empirically-calibrated simulation. The model is concerned with two classes of micro-level mechanisms. Forces of attraction are social mechanisms that explain the creation of positive ties within groups, thereby enforcing group cohesion. Forces of repulsion are social mechanisms that explain why positive ties between groups are less likely to be created and why negative ties are more likely instead, thereby enforcing group boundaries. Fig. 1 illustrates how the central concepts of this paper (group boundaries, group cohesion, forces of attraction, forces of repulsion) are linked to the formation of structural groups.

Models explaining groups with forces of attraction and repulsion include the social distance model of Blau (1994), the segregation model of Schelling (1978), and force-directed models applied in graph drawing (Brandes, 2014). Further, Skvoretz (2013) proposes a model of in-group attraction and out-group repulsion within larger societies. Grow and Flache (2011) discuss the emergence of groups along demographic faultlines. These models have in common that group boundaries are shaped by exogenous attributes and distances. It is possible, however, that group boundaries are independent of such exogenous factors and evolve through multiple endogenous social network mechanisms. Human groups do in fact not only occur along pre-existing attributes, distances, or kinship ties. Even randomly assembled groups have been linked to in-group attraction and between-group repulsion (e.g., Sherif et al., 1961; Puurtinen et al., 2015). One approach of modeling the emergence of structural groups is to assume that the cost of individual ties are non-linear (e.g., individuals maintaining a fixed number of N ties) and that thereby groups with a group size of N are an optimal outcome for individuals that simultaneously strive for being embedded in embedded structures. If we refrain from making such assumptions, however, it is less clear how to model an endogenous process of group formation.

The model that we propose builds upon social network theory and incorporates a battery of well-known positive social mechanisms, for example, reciprocity, transitivity and degree popularity (Rivera et al., 2010; Kadushin, 2012). These mechanisms can be linked to the creation of positive ties within an existing group (group cohesion) and can operate as forces of attraction. However, in the absence of exogenous variables these forces of attraction cannot explain why individuals refrain from creating ties to out-group members. Therefore, we further consider social mechanisms on the evolution of negative ties (Labianca and Brass, 2006; Labianca, 2014; Pál et al., 2016). For example, building on arguments from balance theory and identity theory, we can express how individuals within the same group agree on their negative perceptions of out-group individuals, rather than forming positive ties with them. We argue that the incorporation of such forces of repulsion into a network model helps to explain the emergence and stability of group boundaries.

We discuss the effects of various social network mechanisms (micro-level) on the emergence and stability of groups (macro-level) in section 2. We develop a theoretical model which incorporates both forces of attraction and repulsion. We compare this to a model in which only forces of attraction are considered. We anticipate that the model which includes forces of repulsion will be better able to explain the emergence and stability of structural groups. We test this proposition empirically in a two-step procedure, using a combination of two established analytical strategies.

In the first step, we fit stochastic actor-oriented models (SAOMs, Snijders, 2017) to empirical network data of 479 secondary school students from 13 school classes (between whom 7,879 friendship ties and 3,056 negative ties are observed). We estimate two micro-level models that take into account a number of mechanisms of attraction (attraction model) and mechanisms of attraction and repulsion (attraction and repulsion model). The empirical setting is described in section 3. In order to derive a parsimonious SAOM, we develop a set of novel specifications that express the co-evolution of positive and negative networks. The results of the SAOMs are discussed in section 4.

In the second step, we treat the estimated model as an empirically calibrated agent-based simulation model (Macy and Willer, 2002; Bianchi and Squazzoni, 2015; Snijders and Steglich, 2015; Stadtfeld, 2018). We investigate whether the two empirical models are able to generate networks in which structural groups emerge (section 5). In accordance with our theoretical expectations, we find that the model with only forces of attraction is not able to generate groups with clear boundaries that are stable through time. This macro-outcome is better represented when forces of attraction and repulsion are considered simultaneously. In section 6, we reflect on our combination of research strategies from agent-based modeling and statistical network analysis and discuss in how far the SAOM is different from traditional agent-based simulation models. In section 7, we discuss the implications of our model, results, and analytical strategy for future research on explaining groups and other macro-level phenomena in social networks.

2. A theoretical model of group formation and stability

We develop a theoretical micro-macro model of group formation. The model builds upon existing social network theory and links a variety of micro-level network processes to group cohesion and group boundaries. It illustrates how the emergence and stability of groups in social networks (the macro level) can be explained by a combination of forces of attraction and repulsion.

2.1. Micro level: Forces of attraction and repulsion

Individuals form positive social ties for various reasons (see for example, Rivera et al., 2010; Schaefer et al., 2010; Kadushin, 2012; Block, 2018). The four most widely discussed positive mechanisms in directed social networks are sketched in the first row of Table 1. These “big four” social network mechanisms are reciprocity, transitivity,
advantage), and homophily3. These positive network mechanisms are related to a number of negative network mechanisms that are shown below in rows 2–5 of Table 1. We review and further develop a number of arguments that address why these mechanisms tend to be dominant dynamic patterns in social networks (see Yap and Harrigan, 2015), and how they are related to group cohesion and group boundaries.

Reciprocity (Table 1, first column), or the mutual alignment of sentiments and perceptions in dyads, is a process operating between pairs of actors. Granovetter (1973) argues that strong social ties are characterized by “mutual confiding” and “reciprocal services” (p. 1361). Other mechanisms behind reciprocity explain how one-sided relationships tend to either dissolve or develop into mutual ties (Hallinan, 1978). Reciprocity may explain how cohesion between two individuals increases through time and, if one-sided relationships are more likely to emerge within existing groups, how the tendency towards reciprocity will further promote group cohesion.

Below positive reciprocity in Table 1, we show three related dyadic mechanisms. Reciprocity in negative ties represents a tendency for mutually shared negative perceptions between two actors. This mechanism is relevant in outlining boundaries between groups. The two types of mixed (positive-negative) reciprocity are assumed to be outcomes rarely observed due to tendencies for reciprocity in positive and negative networks and for the dissolution of asymmetric positive ties. Processes that prevent the reciprocation of positive ties with negative ones help explaining how within-group negativity may give way to increased cohesion and how negative group boundaries become clear and stabilize over time.

Transitivity (Table 1, second column) is the tendency for positive ties to be aligned between three actors. There are a number of reasons why positive ties between A and B, and B and C lead to positive ties between A and C. Transitivity may stem from an overlap in interaction opportunities (Festinger et al., 1974), dependence between distances in physical space or Blau space (Blau, 1994), or overlap in social foci (Feld, 1981). It can further be a side-product of shared (and potentially unobserved) homophilous preferences (Granovetter, 1973; Goodreau et al., 2009; Stadtfeld and Pentland, 2015), be a consequence of tie formation in line with balance theory (Heider, 1958; Cartwright and Harary, 1956), or be related to reputation diffusion and hierarchy formation (Chase, 1980). The consequence of transitivity (and similarly, potential tendencies of individuals to close cyclic triadic structures) is increasing cohesion within subsets of the network where positive relationships already exist. Transitivity may serve as a nucleus of group formation and thus operates as a central force of attraction. Mechanisms of transitivity and clustering that interplay with reciprocity can magnify these effects (Davis, 1970).

The emergence of transitivity through mechanisms of cognitive balance has been generalized to triadic structures in which positive and negative ties co-exist (Cartwright and Harary, 1956; Hummon and Doreian, 2003). In this approach, any triadic configuration with three positive relationships or one positive and two negative relationships can be seen as balanced. However, the definition of a positive or a negative relationship is not self-evident in empirical studies, since individuals in a triad may have incongruent perceptions about their ties to one another.4 One could argue that many specific network mechanisms have to be distinguished, for example, those in which a negative tie is asymmetric in either direction or reciprocated.

Here, we propose a simplified representation that significantly reduces the set of possible triadic mechanisms. We assume that a dyad in which at least one negative tie exists is likely to be perceived as a “conflicted dyad” by both individuals involved. This definition is in line with “negative relationships” elaborated in much detail in the social ledger theory of Labianca and Brass (2006). We expect that conflicted dyads are potentially visible to third individuals, affecting the way they form triadic relationships with them. We further assume that an individual A who is not positively tied to either B or C is most likely to perceive that B and C are in a positive relationship if their relation is mutual. We thereby define three mixed (positive-negative) balanced network mechanisms in directed networks (Table 1, second column): i) individuals tend to form positive ties to others if there is a third person with whom they both have a conflicted dyad; ii) individuals tend to form negative ties to others who are in a conflicted dyad with someone who they are positively tied to; iii) individuals tend to form negative ties to others who are reciprocally positively connected with someone with whom they are in a conflicted dyad. The advantage of using conflicted dyads is that we do not need to consider different cases that have the same consequences according to structural balance theory (e.g., disliking someone a friend dislikes and disliking someone who dislikes a friend).

Each of the mixed (positive-negative) balanced triads includes one positive relationship and two negative ones. When groups exist, we may assume that the positive tie is most likely to be within a group, whereas the two negative relationships will define the boundaries to a third person who is not a member of the group (Hummon and Doreian, 2003). The mixed balanced triadic mechanisms thus contribute to group cohesion and group boundaries and operate as forces of attraction and repulsion. This interpretation of balanced triads is in line with social identity theory (Tajfel and Turner, 1979) that argues that people within a group tend to agree on their perceptions towards outgroup members. These shared perceptions will be observed at the groups’ boundaries. Shared perceptions and group boundaries will further contribute to the development of a group identity that enforces group cohesion through positive ties.

Popularity (Table 1, third column) is an endogenous social mechanism and relates to the tendency that those who are perceived as popular by others tend to attract additional positive ties. Merton (1968) initially referred to this phenomenon as the “Matthew effect”, whereas other scholars called this mechanism “preferential attachment” (Newman, 2001), or “cumulative advantage” (Price, 1976). The same

---

3 Block (2018) discusses the co-evolution of three of these mechanisms: reciprocity, transitivity, and homophily.

4 The fact that reciprocity generates tendencies for such incongruent perceptions to be resolved in dyads does not solve this empirical issue, because dyadic and triadic network processes operate simultaneously.
mechanism for negative networks is equally likely: negatively perceived individuals tend to attract additional negative ties (Pál et al., 2016). Positive and negative popularity effects should partly be irrespective of group boundaries and it is thus unclear how much popularity mechanisms in negative or positive networks may contribute to cohesion within groups or to the emergence of group boundaries. However, mechanisms of popularity (as well as mixed popularity effects that are referring, for example, to the tendency not to dislike popular individuals) may explain the formation of initial positive and negative ties that serve as nuclei in the formation of groups and the definition of boundaries.

Homophily (Table 1, last column) relates to the increased likelihood of forming ties to others who are similar (McPherson et al., 2001). Empirical evidence for the effect of homophily on group formation has been found within various contexts (Mäs et al., 2010) and regarding different attributes (see e.g., Moody (2001), Stark and Flache (2012), Boda and Néray (2015) on ethnic segregation). Homophily will increase the cohesion within groups whose members are characterized by an attribute and thus operate as a straightforward force of attraction.

In turn, heterophobia as a parallel mechanism would explain how negative ties generate boundaries between groups of individuals who are dissimilar. In case of the presence of strong homophily and heterophobia, one could thus perfectly explain the emergence and stability of social groups (Skvoretz, 2013). However, homophily has been argued to be multidimensional (Block and Grund, 2014). This would suggest that homophily as a main driver of group formation could be limited to contexts in which multiple individual attributes are highly correlated. Rarely are social groups perfectly bounded by individual attributes which suggests that similarity mechanisms need to be studied in a multi-mechanistic network framework. Some studies (e.g., Marques et al., 1988; Pál et al., 2016) found empirical evidence for mechanisms of dislike between similar individuals (homophobia5). In such situations a homophily/heterophobia model will in particular not be able to explain the formation of social groups. Overall, we assume that similarity-based mechanisms play an important role in promoting group cohesion and drawing group boundaries.

2.2. Macro level: The emergence and stability of groups

Social forces of attraction can naturally be linked to the development of cohesion within groups. Reciprocity explains how one-sided relationships develop into mutual relations. Transitivity explains how “missing links” in a group may come about. Popular individuals within a group will attract additional ties from group members. Homophily explains how those who are similar are more likely to form cohesive groups.

Conceptually, however, there is a flaw in the attraction model on the macro level. If reciprocity, transitivity, popularity, and homophily explain tie formation between individuals, then this happens partly irrespective of endogenous group boundaries. A popular actor within a group will also attract ties from the outside, giving rise to opportunities for subsequent transitive closure and reciprocity. Any incidental connection from a group member to the outgroup serves as a seed facilitating the further growth of the group and the integration of outsiders and their social ties. A pure attraction model will thus, intuitively, only explain how groups expand, until they reach the limit set by the social context. The level of cohesion in this scenario is potentially constrained by the finite time and resources individuals can invest in forming and maintaining relationships, but not by group boundaries. As a consequence, boundaries between groups will seemingly exist at times, but they will be unstable, unless exogenous attributes in combination with homophily can fully explain them. If this is not the case, each pair of individuals will be likely to appear in the same group at some point. This macro-level outcome is contrary to our intuition and empirical observation of stable group structures in various social contexts that do not solely follow exogenously defined faultlines.

Including social forces of repulsion is a possible way of solving the problem of unclear group boundaries and group instability. Then, through mechanisms of structural balance, individuals will not only align their perceptions about positive ties, but also about negative ties (or neutral or weak ties). For example, if two friends A and B dislike individual C, then also C’s friends D and F will tend to dislike A and B. Individuals D and F then not only have a common friend C, but also two common foes A and B which will make them more likely to be friends with one another. Forces of attraction are thus not only dependent on forces of repulsion, but their dynamic interplay will lead to the emergence of coalitions with clear boundaries (A, B on one side; C, D, E on the other). Mixed-network reciprocity and popularity mechanisms will further contribute to a compartmentalization of negative and positive ties within different parts of the network.

Fig. 2 illustrates two prototypical outcomes that we expect from the two models. The network on the left is an expected outcome of the attraction model. It consists of a big main component in which most individuals are indirectly connected. Colors indicate a potential structural identification of groups (e.g., using the approach by Newman, 2006). However, the boundaries between them are unclear and are likely to change through time. The network on the right is an expected outcome of the attraction and repulsion model. Group structures are clearly identifiable in the positive network (solid lines are positive ties) and are expected to be more stable through time, since the repulsive forces stabilize boundaries which in turn further increase group cohesion (dashed lines are negative ties).

2.3. Two propositions

We derive two network models from the discussion above. One model (the attraction model) is concerned with the dynamics of positive social network mechanisms and thus only considers forces of attraction6. The other model (the attraction and repulsion model) is concerned with the joint dynamics of positive and negative social network mechanisms and thus considers forces of attraction and repulsion. We derive two propositions from the discussion.

Proposition 1. Group emergence proposition: In comparison to the attraction model, the attraction and repulsion model generates structural groups with higher levels of group cohesion and clearer group boundaries.

Proposition 2. Group stability proposition: In comparison to the attraction model, the attraction and repulsion model generates structural groups that are more stable over time.

3. Empirical setting: Friendship and dislike in secondary school classes

3.1. Data

We test the propositions about the two models in an empirical setting of 13 secondary-school classes in Hungary. Attraction mechanisms are modeled through friendship nominations; forces of repulsion are modeled through expressed dislike. Friendship and dislike nominations are measured on the same 5-point that ranged from dislike over neutral to friendship, hence there are no ambivalent nominations in the data. School classes are an ideal setting for dynamic social network studies as students spend most of their time within this context, they have well

5 In other contexts, the term homophobia is used for negative attitudes towards homosexuality. This is different from the use in this article.

6 Effects that limit the number of ties that an individual maintains may be interpreted as a general force of repulsion here.
defined boundaries, and they are stable during secondary-school years. We utilize a unique network panel from the study titled “Wired into Each Other: Network Dynamics of Adolescents in the Light of Status Competition, School Performance, Exclusion and Integration” (Boda and Néray, 2015; Vörös and Snijders, 2017; Pál et al., 2016). The data contain four survey waves in the period of 2010–2013. We analyze only the first 3 waves as wave 4 had significantly higher non-response. The first wave took place approximately two months after students first met in school, the second wave followed half a year later, and a third approximately one year after the second. Hence, students were 9th-graders in the first and in the second wave, and 10th-graders in the third wave. Of the 44 classrooms in the sample, only 13 were used for the analysis; see Appendix A for details. Surveys were completed during regular classes in the presence of trained interviewers and took no more than 45 minutes. Students without parental consent were not included in the analysis, and those who were absent during the data collection were coded as missing. Students were assured that their answers would be kept confidential and used for research purposes only.

3.2. Descriptives

Our sample consists of 479 individuals of which 292 are females (61%). Across the three data collection waves we observe 7,879 friendship nominations and 3,056 dislike nominations.

Table 2 shows the 13 empirically observed networks in the second wave. In each panel of the table we show the friendship network on the left in which automatically detected structural groups are highlighted\(^7\). On the right side of each panel we plot the positive and negative relationships together. It appears that the networks mostly have a clear group structure: the percentage of friendship ties within structural groups (identified through modularity maximization; Newman, 2006) ranges from 58% to 92%. It further seems that negative relationships tend to connect individuals who are in different groups of the friendship network (with the rather unstructured network 6 as a potential exception). The percentage of dislike relationships that connect individuals who were identified as being in different groups ranges from 63% to 97%. To compare those numbers to a meaningful reference category, we permute (rewire) the dislike nominations of all individuals to random others and count the number of cases where the percentage of between-group ties is higher than observed in 1000 simulations. By rewiring ties, we maintain the network density as well as the out-degree of each individual. The calculated percentage constitutes a non-parametric p-value. In ten of the 13 networks, the proportion of between-group dislike is higher than expected (86-94%) with p-values between 0 and 0.034. Exceptions are network 3 (89%; \(p = 0.17\)), network 6 (65%; \(p = 0.80\)) and network 12 (74%; \(p = 0.54\)) where we find no evidence for a higher proportion of between-group dislike than expected. These findings are static and descriptive but relate to Proposition 1 on the emergence of groups.

Table 3 shows the 13 networks in wave 1 and 2. The right network in each panel (wave 2) corresponds to the respective classroom network in Table 2. On each network the same modularity-optimizing algorithm used earlier was run; structural groups are marked by highlighted areas. Node colors in both waves represent the structural groups of the nodes in the first-wave network. One can see that a high proportion of groups seem to be relatively stable through time as only few nodes “change” (for example, an orange node appears in the green group in wave 2 of network 1). Some group partitions seem very stable (e.g., networks 1 and 10), others somewhat less (networks 7 and 8). The proportion of pairs of individuals that are in the same structural group across the two waves ranges from 64% to 92%. These values are high given that the stability of the friendship network is relatively low: the proportion of

---

\(^7\) Groups were identified using an edge betweenness clustering algorithm that maximizes the modularity criterion on a symmetrized transformation of the friendship network (Newman, 2006). Within each dyad, the maximum of both ties was used for the symmetrization. Networks are visualized with the igraph software in R (Kolaczyk and Csárdi, 2014).
ties that are present in both waves among ties that are present in either wave (the Jaccard coefficient) only ranges from 26% (network 12) to 55% (network 1)\(^8\). We assess the significance of the group stability wave (the Jaccard coefficient) only ranges from 26% (network 12) to 55% (network 1)\(^8\). Structural groups in the friendship network are highlighted. The node color corresponds to the groups in wave 1. Isolated nodes in the friendship networks are not shown.

| 1 | 2 |
|---|---|
| 3 | 4 |
| 5 | 6 |
| 7 | 8 |
| 9 | 10 |
| 11 | 12 |
| 13 |  |

---

Table 3
Empirical friendship networks in the first and the second wave in the 13 classrooms. In each panel, the network in wave 1 is plotted on the left and the network in wave 2 on the right. Structural groups in the friendship network are highlighted. The node color corresponds to the groups in wave 1. Isolated nodes in the friendship networks are not shown.

4. Evidence for micro-level network processes of attraction and repulsion – a SIENA model

We fit two multigroup stochastic actor-oriented models (SAOMs, Snijders, 2017) to the longitudinal network data described in section 3. Models are estimated with the RSiena software (Ripley et al., 2019)\(^9\). One model expresses positive social mechanisms of attraction (attraction model) and one model expresses the simultaneous operation of positive and negative social mechanisms (attraction and repulsion model). The two submodels of the second model (positive and negative network dynamics) are treated as interdependent processes. Model parameters are estimated with the method of moments (Snijders, 2017). It ensures that for each effect included the model (e.g., reciprocity), a simulation from the model will return a network in which the corresponding network-level statistic (e.g., the number of reciprocal ties) across all groups is expected to be equal to the empirically observed statistic in the subsequent wave.

The models are specified by a number of effects that are representative of micro-level social mechanisms that involve two or a few individuals. They express, for example, whether individuals are more likely to dislike those who their friends dislike. Given the theoretical complexity of triadic social mechanisms that include both negative and positive ties, a set of new RSiena effects was developed for this article. These new effects relate to the specification of balanced triadic mechanisms in directed networks. Details are presented in Appendix C.

Results of the first model (attraction model) are shown in Table 4. Results of the second model (attraction and repulsion model) are presented in two tables: Table 5 includes the results of the submodel about positive network dynamics, Table 6 includes the results of the submodel about negative network dynamics. Both submodels include parameters that express how the two networks co-evolve.

The first 13 effects of the attraction and repulsion model in Table 5 are equivalent to the effects in the attraction model in Table 4. The estimated parameters of these effects are similar in both models (the reciprocity parameter, for example, is 1.97 and 1.86, respectively). We thus focus on the interpretation of the results of the attraction and repulsion model in Tables 5 and 6 in the following. In the interpretation of effects we follow the structure outlined in Table 1 and discuss in the following subsections reciprocal, triadic, degree-related, and gender-related mechanisms separately. One additional parameter of the model are the “rates” that explain how often actors in the network consider changing their outgoing ties between two subsequently observed networks. The mean rates range from 8.9 friendship network changes per actor (attraction model) to 12.3 negative tie changes per actor (attraction and repulsion model) within an observed period. Another central parameter is the “density” that express that the number of ties that actors typically maintain is limited (e.g., due to cognitive or time constraints).

Dislike and friendship co-evolve. They are modeled as disjoint categories, in line with the questionnaire design in which those were assessed with a single scale. The SAOM is therefore specified similarly to an “ordered SAOM” (e.g., Elmer et al., 2017). An individual A can only establish a friendship nomination to B if A does not dislike B. The model assumes that individuals who want to switch a tie from friendship to dislike first have to dissolve the friendship tie.

All micro-level mechanisms that are included in the models are visually represented in Table 1 (core mechanisms) as well as in Table 9 in Appendix C (additional control mechanisms).

4.1. Dyadic network mechanisms

Effects 3 and 19 in Tables 5 and 6 provide evidence that individuals tend to reciprocate both positive and negative ties. Effects 14 and 29 further weakly indicate the tendency not to reciprocate positive or negative ties with a tie of the opposite valence (an opposite tie creation and maintenance is less likely). Those two effects are negative, but are

\(^8\)The Jaccard coefficient ranges from 0 to 1. An index of 1 indicates that the network does not change between two waves, an index of 0 indicates that all ties are only present in either of the two waves but not in both.

\(^9\)Details on the multigroup option in RSiena are discussed in Ripley et al. (2019, section 11.1). Appendix B provides more details.
Table 4
Attraction model: Estimated effects of positive micro-level mechanisms

| dep | name            | ref    | fig | est  | se   | sig |
|-----|----------------|--------|-----|------|------|-----|
| 1.0 | positive mean rates |        |     | 8.93 | 3.88 |     |
| 2.0 | positive density  | Fig. 6a|     | -1.23| 0.07 | *** |
| 3.0 | positive recip    | Fig. 6b|     | 1.97 | 0.07 | *** |
| 4.0 | positive transTrip | Fig. 6c|     | 0.31 | 0.02 | *** |
| 5.0 | positive transRecTrip | Fig. 6d|     | -0.15| 0.02 | *** |
| 6.0 | positive inPop    | Fig. 6e|     | 0.02 | 0.01 | *   |
| 7.0 | positive outPop   | Fig. 6f|     | -0.15| 0.01 | *** |
| 8.0 | positive outAct   | Fig. 6g|     | -0.03| 0.00 | *** |
| 9.0 | positive outTrunc | Fig. 6h|     | -1.29| 0.29 | *** |
| 10.0| positive antiInIso| Fig. 6i|     | -0.86| 0.26 | **  |
| 11.0| positive altX     | Fig. 6j|     | -0.07| 0.04 | *   |
| 12.0| positive egoX     | Fig. 6k|     | 0.00 | 0.04 |     |
| 13.0| positive sameX    | Fig. 6l|     | 0.23 | 0.04 | *** |

The columns indicate the dependent variable (here: change in the positive network), the technical RSiEa short name, a reference of the mechanism in the figures illustrating the model specifications, a visual representation of the mechanism, estimates, standard errors (standard deviations in case of the mean rates), and levels of confidence.

\*p < 0.1.
\**p < 0.05.
\***p < 0.01.
\****p < 0.001.

Table 5
Attraction and repulsion model (part 1: dependent variable is change in the positive network)

| dep | name            | ref    | fig | est  | se   | sig |
|-----|----------------|--------|-----|------|------|-----|
| 1.0 | negative mean rates |        |     | 10.89| 3.72 |     |
| 2.0 | negative density  | Fig. 6a|     | -1.18| 0.08 | *** |
| 3.0 | negative recip    | Fig. 6b|     | 1.86 | 0.07 | *** |
| 4.0 | negative transTrip | Fig. 6c|     | 0.29 | 0.02 | *** |
| 5.0 | negative transRecTrip | Fig. 6d|     | -0.14| 0.02 | *** |
| 6.0 | negative inPop    | Fig. 6e|     | 0.01 | 0.01 |     |
| 7.0 | negative outPop   | Fig. 6f|     | -0.15| 0.01 | *** |
| 8.0 | negative outAct   | Fig. 6g|     | -0.02| 0.00 | *** |
| 9.0 | negative outTrunc | Fig. 6h|     | -1.28| 0.27 | *** |
| 10.0| negative antiInIso| Fig. 6i|     | -0.61| 0.27 |     |
| 11.0| negative altX     | Fig. 6j|     | -0.06| 0.04 | *   |
| 12.0| negative egoX     | Fig. 6k|     | -0.03| 0.04 |     |
| 13.0| negative sameX    | Fig. 6l|     | 0.27 | 0.03 | *** |
| 14.0| negative crprodRecip| Fig. 6m|     | -0.58| 0.32 |     |
| 15.0| negative inPopIntn | Fig. 6n|     | -0.05| 0.01 | *** |
| 16.0| negative WWX.EE   | Fig. 6o|     | 0.06 | 0.01 | *** |

The columns indicate the dependent variable (here: change in the positive network), the technical RSiEa short name, a reference of the mechanism in the figures illustrating the model specifications, a visual representation of the mechanism, estimates, standard errors (standard deviations in case of the mean rates), and levels of confidence.

\*p < 0.1.
\**p < 0.05.
\***p < 0.01.
\****p < 0.001.

Non-significant effects capture potentially relevant information that can be exploited in a simulation study; we argue that their inclusion is preferred over dropping non-significant effects and thereby assuming a zero effect. We will later use non-significant estimates, like the negative mixed reciprocity effect, in the agent-based simulations.
4.3. Degree-related network mechanisms

Ten social mechanisms relate to the degree distributions in the positive and negative networks. Effects 6-10 in Table 5 and effects 20-24 in Table 6 capture degrees dynamics. Mechanisms 6 and 20 are of specific interest as they relate to preferential attachment in the positive and negative network as shown in Table 1. We find no evidence for in-degree popularity or preferential attachment (Merton, 1968; Price, 1976) in the friendship network (the effect is positive but not significant). However, we find evidence for in-degree popularity in the dislike network – Pál et al. (2016) refer to this mechanism as to the “black sheep” effect. In the negative network we find a tendency towards outdegree dispersion with few individuals disliking many others (effect 22), whereas in the friendship network outdegrees seem to be more homogenous (effect 30).

Both networks show a strong tendency for a few individuals not to have any ties and thus not to participate in the friendship or disliking structure of the class (effect 9 and 23). In both networks we observe a tendency not to nominate those who call many others their friends (effect 7) or who dislike many (effect 21) which could indicate a hierarchical structuring of the two networks. In both networks it is unlikely that individuals who have an in-degree of zero (are not called a friend or disliked by anyone) subsequently attract friendship nominations (effect 10) or dislike (effect 24).

The mixed-degree effects 15 and 30 in Tables 5 and 6 relate to how degree effects operate across networks. We find evidence that individuals tend not to become friends with others who are disliked by many (effect 15), but not for the opposite mechanism that individuals tend not to dislike others with many friends (effect 30).

Table 7
The first 9 network simulations of friendship networks from the attraction model in Table 4. Snapshots are plotted after simulations with rates 10 (close to the empirical rate), 50, 100, 250, 500, and 1000. Clusters with at least two nodes are highlighted by background and node colors.

| rate10 | rate50 | rate100 | rate250 | rate500 | rate1000 |
|-------|-------|--------|--------|--------|---------|
| 1     |       |        |        |        |         |
| 2     |       |        |        |        |         |
| 3     |       |        |        |        |         |
| 4     |       |        |        |        |         |
| 5     |       |        |        |        |         |
| 6     |       |        |        |        |         |
| 7     |       |        |        |        |         |
| 8     |       |        |        |        |         |
| 9     |       |        |        |        |         |

4.4. Similarity mechanisms

Three gender-related social mechanisms were included in each of the submodels; one relates to similarity and the other two to the differential activity (ego effect) and attractiveness (alter effect) of females as compared to males. We find gender homophily in the friendship network (effect 13 in Table 5) but also a tendency that individuals of the same gender are more likely to dislike one another (homophobia; effect 28 in Table 6). Alter effects indicate that female students (coded as 1, males as 0) are slightly less likely to be nominated as friends but more likely to be disliked (effects 11 and 12 in Tables 5 and 6). The strong homophily effect 13 in the attraction model indicates that group cohesion could partly be explained by similarity. Within the attraction and repulsion model, however, it seems unclear however how the gender attribute could contribute to the emergence of groups, given that we also find evidence for disliking others who have the same gender.

5. The emergence and stability of groups on the macro level – an agent-based model

5.1. Exploring the micro-macro link

Through simulations, we now investigate whether the estimated SAOMs can explain the emergence and stability of structural groups (the propositions in section 2.3). To that end, we compare the empirical attraction model to the empirical attraction and repulsion model. We do not aim at explaining which model is better in terms of fit to the data, but to understand in how far the theoretical prediction holds that the attraction and repulsion model will generate outcomes with clearer group boundaries and a higher level of stability. The answer to this question is not self-evident, as the models are constructed from micro-level mechanisms that are mostly concerned with two (dyadic and similarity mechanisms) or three (triadic mechanisms) actors. Structural groups, however, typically consist of more than three nodes and are simultaneously characterized by group cohesion and group boundaries. We utilize the fact that SAOMs express actor-oriented micro-level processes and that these processes can be straightforwardly simulated (as illustrated by Snijders and Steglich, 2015; Prell and Lo, 2016; Block et al., 2018; Stadtfeld, 2018; Stadtfeld et al., 2018). Simulations are carried out with the RSiena software (Ripley et al., 2019). Applying SAOMs as agent-based simulation models is related to a prominent literature strand about agent-based models (ABMs) in sociology (e.g. Macy and Willer, 2002; Gilbert, 2008; Macy and Flache, 2009; Bianchi and Squazzoni, 2015).

From each of the two models (attraction model, attraction and repulsion model), we simulate 100 long chains of evolving networks with a maximum rate of 1000 (about 100 times the empirically estimated rate). The simulations start from an empty network with 36 nodes (the average size of the empirically observed networks). The same analyses were conducted on networks with 200 nodes which essentially returned the same results. Networks can be drawn from the simulation chain at any point in time.

We show nine exemplary simulation chains of the attraction model (Table 7) and nine exemplary simulation chains of the attraction and repulsion model (Table 8). Each simulation is illustrated in one row of the tables. Networks are drawn from the simulation chain after a rate of 10, 50, 100, 500, and 1000 (the columns of the tables). A rate of 10, for

11 SAOMs model network change as a result of individuals making choices about their set of outgoing ties.

12 The mean estimates of the rates are 8.9, 10.9 (friendship) and 12.3 (dislike) and are shown as effect 1 in Table 4 (attraction model) and effects 1 and 17 in Table 5 and 6 (attraction and repulsion model). For more details about the rate parameters in estimation and simulation with SAOMs, see Ripley et al. (2019).
Table 8
The first 9 network simulations of friendship networks from the attraction and repulsion model in Tables 5 and 6. Snapshots of the friendship ties are plotted after simulations with rates 10 (close to the empirical rate), 50, 100, 250, 500, and 1000. Dislike ties are omitted from the plot, but not from the simulation model. Clusters with at least two nodes are highlighted by background and node colors.

| rate10 | rate50 | rate100 | rate250 | rate500 | rate1000 |
|--------|--------|---------|---------|---------|----------|
| ![Network Simulation](image) | ![Network Simulation](image) | ![Network Simulation](image) | ![Network Simulation](image) | ![Network Simulation](image) | ![Network Simulation](image) |

Table 9
Additional social mechanisms that relate to attraction (+) and repulsion (–).

| Density | Triadic | Two-Paths | Out-degree |
|---------|--------|-----------|------------|
| (+)     | ![Mechanism](image) | ![Mechanism](image) | ![Mechanism](image) |
| (–)     | ![Mechanism](image) | ![Mechanism](image) | ![Mechanism](image) |

example, indicates that in the simulation each actor reconsiders her set of outgoing ties on average 10 times, starting from an empty network. The larger the rates, the less meaningful is an interpretation of rates in terms of actors’ opportunities to reconsider their set of ties—large rates of, for example, 500 are best interpreted as approaching a stochastic equilibrium of the actor-oriented network change process. In a stochastic equilibrium, the macro-level features of the process are relatively stable while change still occurs on the micro-level. The rates were chosen so that different phases of the simulation process (as discussed below) are represented in the visualizations. To visually capture the formation of group structure, only the friendship networks are plotted in the two tables even though in the attraction and repulsion model in Table 8, a co-evolving dislike network was part of the simulation.

The visual representations are in line with the micro-macro link that we established in the theoretical model. In the attraction model in Table 7, we cannot see that groups emerge within the simulated time span. Groups tend to be overlapping without clear boundaries between them. It is noteworthy that this is the fact even though the gender homophily mechanism could generate group structures in principle. In the attraction and repulsion model in Table 8, in contrast, we can see that with rates of 250 or more group-like structures emerge that have a high cohesion within and no or few connections between them. Isolated nodes are, however, more frequent in the attraction and repulsion model. The reason for this phenomenon is that social mechanisms that explain the formation of group boundaries (e.g., agreement on disliked individuals) can also explain emergence of isolated nodes in a network. Over time, attraction forces within groups tend to become so strong that the within-group friendship networks get close to complete subnetworks in some cases. This observation is related to model degeneracy discussed in the networks statistics literature (e.g., Snijders, 2002).

Systematic comparisons of features of 100 simulated network chains per model are presented in Figs. 3–5. Here, we aggregate descriptive results from the network chains throughout the simulation process. In contrast to the visual inspection above, we now simulate from four empirical observed networks of the first data collection wave and compare the resulting statistics to those of the networks in the second wave (indicated by horizontal lines). As empirical starting points, we chose the wave 1 networks from the first four panels of Table 3. The four empirical cases are shown in the four rows of Figs. 3–5. Statistics refer to the network of positive ties and are shown for the attraction model (blue line) and attraction and repulsion model (red line). All panels further indicate the first and third quartile of the empirical distribution of statistics. Vertical lines indicate rates of 10, 50, 100, 250, and 500 that are in line with the visualizations in Tables 7 and 8.

Fig. 3 relates to proposition 1 (group emergence proposition) and shows the percentage of ties within structural groups (clusters). The three columns present the statistic up to a simulated rate of 10 (roughly the empirical rate, see Tables 4–6), 100, and 1000. Within a rate of 10, the simulations develop similarly in terms of within-group ties. Within a period of 100, however, they start developing different patterns. In three out of four cases, the clustering generated by the attraction and repulsion model increases and eventually transitions to a stochastic equilibrium in which almost all ties are within clusters. From the visual inspection, we could conclude that in the long run the attraction and repulsion model tends to generate networks with completely separated components. In one out of four cases (where the empirical rate of within-group ties is low), the two models keep behaving similarly with respect to the first statistic.

Fig. 4 relates to proposition 2 (group stability proposition) and shows the percentage of within-group ties that are stable in a forward simulation after a rate of 10. The left column includes forward simulations up to a rate of 50. The group stability is evaluated for the networks simulated after a rate of 10. Both models behave similarly at the beginning. After a forward simulation of about 50 (column 2) the stability of the attraction and repulsion model increases in comparison to the attraction model in three out of four cases. In the long run, the structural groups generated by the attraction and repulsion model are extremely stable.

Fig. 5 presents three additional statistics: the percentage of homophilous ties, the number of detected groups (ignoring isolates), and the density of the friendship network. Again, the simulations from the first three empirical networks behave similarly. First, it can be seen that the clustering of the attraction and repulsion model is explained less by the gender attribute and that the percentage of homophilous ties is thus generally lower. Second, the number of clusters in the attraction model

---

13 This definition of equilibrium thus differs from static equilibrium concepts known in ABM applications that are derived from game-theoretic models.
is typically higher at a value of 5–7 in comparison to 2–4 clusters in the attraction and repulsion model. Third, the density of attraction and repulsion model is higher in the long run (25–30%) as the density within clusters increases to values close to 1. The fourth empirical case is again different. It can be seen that in case of the attraction model typically only one structural group emerges in the long run and that the network density increases. It seemsthat in this case both models suffer from (local) model degeneracy where extreme stochastic states are reached with very high or low densities. The probability of transitions between these regimes is supposedly very low.

6. Reflection on applying stochastic actor-oriented models as ABMs

This section reflects on a perceived gap between the literature on ABMs and empirical network modeling that researchers may come across when aiming at speaking to the two communities simultaneously. Overall, we believe that the integration of the two approaches is quite natural and promising. Bianchi and Squazzoni (2015, p. 284), for example, define ABMs as “computer simulations of social interaction between heterogeneous agents [...], embedded in social structures [...] that are built to observe and analyze the emergence of aggregate outcomes.” Macy and Willer (2002, p. 146) define ABMs as simulation models of agents that are autonomous and interdependent, that follow simple rules, are adaptive and backward-looking. Gilbert (2008, p. 2) defines agent-based modeling as a “computational method that enables a researcher to create, analyze, and experiment with models composed of agents that interact within an environment.” The purpose of our simulations is indeed the observation and analysis of an emergent aggregate outcome (structural groups). Individuals within the SAOM take autonomous but interdependent decisions on the creation of ties.
(positive and negative) which changes their position and decision opportunities in the network. This interdependence is observed when, for example, a reciprocal tie is generated as a sequence of two tie creations. The newly created reciprocal tie may then create opportunities for further subsequent network changes (e.g., reciprocity → popularity → transitivity) that alter the positions of the individuals in the network, for example, their level of embeddedness. These interdependent tie creations can be understood as a form of interaction. The changing positions of actors as well as their differences in terms of exogenous attributes (gender in our empirical example) induce heterogeneity in this process.

The ABM that we propose in this paper is certainly not prototypical for the ABM literature within sociology. Other than in many ABM studies, for example, we do not explore / experiment with a space of parameters, but limit the parameters to the SAOM parameters estimated from empirical data. Thereby, we are closer to an empirically calibrated ABM as discussed by Boero and Squazzoni (2005), Manzo (2007) and Hedström and Åberg (2005). A parameter exploration only happens in terms of the comparison of two models and by varying a time parameter that governs the length of the simulation process. Our model is more complex than many prototypical ABMs (some famous examples are discussed in Bianchi and Squazzoni, 2015). The attraction and repulsion model, for example, has 32 parameters that aim at taking into account the multi-mechanistic nature of social network dynamics (as discussed in Stadtfeld, 2018). Other differences to classical ABM studies are rooted in assumptions of the SAOM framework. SAOMs do not easily allow the expression of strategic or forward-looking behavior which could be an interesting issue to explore in the context of group emergence. Adaptation of behavior is not taken into account in our simulations, but only occurs due to changes in network position (e.g., after the emergence of groups, in-group preferences are more pronounced).

Fig. 4. Stability of structural groups after a rate of 10 (P2). Median statistics of 100 simulated friendship networks with the attraction model (blue line) and the attraction and repulsion model (red line); the first and third quartiles are indicated by the blue and red areas. The x-axis shows the simulation steps per actor (rate parameter). No values can be calculated for rates smaller than 10. The four rows correspond to four networks (Number 1–4 in Table 3) that the simulations started from. The gray horizontal bar indicates the statistic in the empirical case.
Many ABM studies are further inspired by game-theoretic interactions, which are also different from the SAOM approach.

An important concern in ABMs is empirical calibration. The SAOM is empirically calibrated by means of the parameter estimation: the method of moments estimation ensures that the fit of each included statistic (e.g., reciprocity) is correct in the empirical context and given the estimated parameters. However, once parameters change (the rate is varied in our case) or the empirical context is altered (simulations start from empty networks in our case) this calibration does not hold anymore and it is possible that networks are generated that are different from the empirically observed networks. Indeed, we see that the long-run outcomes of the simulated models are too stylized to be meaningfully compared to empirical data. Yet, we know that the complex set of model parameters are valid in at least one specific empirical context.

We believe that exploring macro-level outcomes based on empirically-calibrated ABMs is a promising research direction (Hedström, 2005). In particular, this is the case because many agent-based models are implicitly or explicitly concerned with group formation and the consequences of in-group and out-group ties on interaction between agents (e.g., cooperation, trust, punishment, see the review of Bianchi and Squazzoni, 2015). Theoretical arguments in this context are concerned with in-group favoritism and negative out-group ties (e.g., Sakoda, 1971; Hammond and Axelrod, 2006). The underlying behavioral model of SAOMs and its ability to be simulated indeed has strong similarities to ABM models\(^{14}\). Snijders and Steglich (2015) proposed first that SAOMs can be applied for this purpose. Stadtfeld

---

\(^{14}\) Empirical SAOM studies have in fact referred to “stochastic agent-based models” to highlight this similarity (e.g. Lomi et al., 2011)
7. Discussion and conclusions

The social lives of humans are organized in small informal groups. These groups are observable in various contexts from schools to digital communities. They are typically characterized by high internal cohesion, observable in tightly knit networks of positive relations, like friendships, connecting group members. The boundaries between groups are in contrast characterized by weaker ties, neutral ties, and sometimes negative ties.

In this paper, we argued that forces of attraction that explain the formation of positive ties within groups of individuals are necessary to express group cohesion, but that they cannot fully explain how stable groups emerge within a larger social context. The reason is that they lack a way of expressing the formation of group boundaries. Those could be expressed by also considering forces of repulsion, that explain the lack of positive ties formation between groups. Based on this intuition, we developed a theoretical framework of group formation in social networks. It builds upon a combination of micro-level social network theories about the evolution of positive and negative relations, and thereby illustrates how forces of attraction and repulsion could jointly explain the emergence and stability of social groups.

We tested the propositions of the theoretical model with a novel combination of dynamic empirical social network analysis and agent-based modeling. As a starting point, we chose an empirical setting of co-evolving networks of friendship and dislike among 479 secondary school students. Descriptively, the networks that they form exhibit group structures, where dislike relations indeed mostly connect individuals in different groups. In a statistical analysis with a newly extended version of stochastic actor-oriented models (SAOMs), we then found evidence for a number of micro-level mechanisms that shape the dynamics of the analyzed networks. Friendship networks, for example, tended to exhibit reciprocity, transitive clustering, gender homophily, and popularity. In the evolution of the dislike networks, we found evidence for reciprocity, disliking of isolated individuals, same gender dislike, and the black sheep effect. The co-evolution of the two networks was characterized by the alignment of friends’ dislike towards other students, and an increased likelihood of individuals to be friends when they dislike the same others.

We then explored how the micro-level mechanisms of two SAOMs (one only modeling forces of attraction, the other forces of attraction and repulsion) were able to generate the emergence and stability of group structures on the macro level. We did so by conducting computer simulations, treating the empirically estimated SAOMs as empirically calibrated agent-based models (ABMs). We could show that in line with our theoretical model, only the model that combined forces of repulsion and forces of attraction generated stable structural groups. In the long run, however, the group features of the simulated networks were too stylized with extremely high densities within groups and a large proportion of isolated nodes. These observations call for further theoretical and statistical work, investigating homogeneity of the micro-level processes contributing to group formation and possible problems of model degeneracy. Comparing the two models, it is noteworthy that even though the attraction model included a relatively strong homophily effect, the generated groups were generally less pronounced and less stable through time.

The current study has general implications for future research on emergent social structures. We presented an extension of the still dominant social network approach, which relies on the analysis of a single relational dimension, by considering the co-evolution of multiplex networks, one positive and one negative. The most straightforward extension by one additional network dimension (negative relations) helped to better explain the emergence and stability of groups on the macro level. The existence of explicit dislike or conflict ties between groups, however, is not a strict necessity for the proposed explanation of group formation. Indeed, one can observe stable groups which are not in direct conflict with one another. In these cases, forces of repulsion can still be present and promote the formation of non-negative, neutral, or weak relationships between groups (rather than stronger ties that tend to be observed within). For example, group norms and identities may exist that simply prohibit group members from interacting and developing positive ties with the outgroup. We encourage future research to test the analytical strategy in different empirical contexts and thereby aim at generalizing the social network theory of group formation that we proposed. Future research could aim at defining more parsimonious variations of the proposed attraction and repulsion model, and evaluate the relative importance of the micro-level mechanisms. Another important extension of this model would be to multidimensional group processes where we do not assume that each individual is the member of one or zero groups. Individuals are in fact often members of multiple groups and these can partly be overlapping. This has not been considered in our “single-dimensional” definition of structural groups.

The presented multiplex network approach may further be used to explain other complex macro-level phenomena. For example, the emergence of social roles within groups or a larger community may be described as the co-evolution of positive ties and perceptions about role-taking behavior (consider the black sheep and popularity effects in our study). Similarly, internal structures of groups such as hierarchies, and their occasional instability and dissolution, may be understood through micro-processes of status competition as expressed by co-evolving positive, negative, and status perception networks. Such efforts may help establish the relation between micro-level processes and the emergence and stability of complex macro-level phenomena in social networks.

Recently, the discussion of group cohesion, group boundaries, and dislike between groups has gained momentum in the context of a perceived increase in polarization in political and societal debates. Particularly, in the context of digital societies there is anecdotal evidence for forces of between-group repulsion that go hand in hand with an increase in within-group cohesion. In the context of digital communities, a lack of interconnectedness between groups may result in “filter bubbles” in terms of information exchange, and in the long run in societally relevant and undesirable outcomes such as segregation, polarization and expressions of hate (e.g., Clemm von Hohenberg et al., 2017). In such circumstances, systematic efforts into explaining the emergence of macro-level structures in human communities are timely. Our theoretical model proposes to consider both forces of attraction and forces of repulsion when trying to understand the dynamics of group emergence and group stability.

Acknowledgments

We thank members of the Social Networks Lab at ETH Zürich and of the social networks research group at the University of Groningen for helpful comments and suggestions that considerably improved this work. The study was supported by the Swiss National Science Foundation (funding no. 10001_A169965). This project has received funding from the European Research Council(ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement no 648693). Data used in this work has been gathered with the support of the Hungarian Research Fund (OTKA, grant number K81336) and the Lendület Grant of the Hungarian Academy of Sciences (K.T.).
Appendix A. The empirical data

A.1 Selection of classes

The complete dataset was collected in seven Hungarian schools and included N = 1,044 students from 44 classes in the first wave with four network waves in total. For computational reasons we reduced the full dataset roughly following the strategy of Pál et al. (2016) and in line with recommendations provided by the RSiena manual (Ripley et al., 2019). The average missing rate was 8% per data collection wave (Details are provided in Pál et al., 2016, p.809). We first dropped the last data collection wave as major composition change happened between the third and the fourth waves that was likely to lead to algorithmic convergence problems (Stadtfeld et al., 2018). From the 40 classes we first selected 19 classes based on a minimum stability criterion between data collection waves. This criterion is proposed by Ripley et al. (2019, p.20) and excludes datasets in which the Jaccard index for two subsequent waves is below a certain threshold, in our case we set it to 0.08. The Jaccard index measures the number of ties that are present in two subsequent waves over the number of ties that are present in either or both of the two waves. A value of 0.08 thus means that 8% of all dislike or friendship nominations have to be stable for the network to be included in the analysis. After the selection of 19 classes, another six school classes were removed from the sample as they lead to problems with convergence of the estimation routine (3 cases), or because parameter estimates would diverge. Reasons for convergence problems can, for example, be caused by missing data and turnover of participants between panels (Stadtfeld et al., 2018). We thus ended up with 13 classes with three data waves each (N = 479). Classrooms in the resulting subsample differed from our total sample along several important dimensions, such as gender composition and type of education. Classes dropped from the analysis turned out to be mainly vocational training and grammar schools with a lower average socio-economic status of students (Pál et al., 2016).

A.2 Measurements

The data of our study was collected by means of self-administered paper-pencil surveys that included social network measurements using a full network roster of the class in several dimensions. Relational information was collected between classmates only. Friendship and disliking ties were self-reported and were part of the same unidimensional five-point Likert-scale. Each student had to indicate his or her relationship with each classmate according to the following descriptions: “I hate him/her” (-2), “I dislike him/her” (-1), “He/she is neutral to me” (0), “I like him/her” (+1), or “He/she is a good friend” (+2). We merged the values -1 and -2 of the scale to create social networks of disliking, given that hate can be conceived of as a strong form of disliking. For friendship, we used the +2 value of the scale. Based on this scale, we created two binary adjacency matrices (disliking and friendship) for each school class in each of the three waves. If student A disliked student B, then the corresponding entry (A,B) in the disliking matrix was marked 1 (0 otherwise). If a student joined the class after the data collection started or he or she left, and would therefore not appear in some of the questionnaires, this student was nevertheless included in all matrices. All values in the corresponding row and column, however, were marked as “structural zeros” to indicate that sending or receiving nominations was impossible in the respective wave (as the student was not the member of the class in the given period). Details about data coding in RSiena are provided in Ripley et al. (2019).

Appendix B. Fitting multilevel stochastic actor-oriented models

We use the RSiena “multigroup” option (Ripley et al., 2019, p.104) to estimate multilevel stochastic actor-oriented models from 13 school classes. Thereby, we assume that the same structural processes govern network change across all classes and that only the change rates may differ. More advanced multilevel analysis techniques like Random Coefficients Multilevel SAOMs (Boda, 2018; Raabe et al., 2019; Ripley et al., 2019, p.105) allow the inclusion of group level variables and to define parameters as random or fixed effects. In our case, however, the multigroup estimation is actually better, as our goal is to determine one set of parameters for our simulation that do not randomly vary between classes but can be used as a global simulation parameter.

Appendix C. Overview of additional effects

Table 9 shows a number of additional social mechanisms that were included in the SAOMs besides the mechanism sketched in Table 1. Density effects in both submodels can be understood as an intercept that assigns a “cost” to the creation of ties and thereby models the overall density of the network. Two additional triadic effects were included. In the positive network we control for the interaction between reciprocity and transitivity, which has been argued to express the relative importance of tie reciprocation in transitive structures (Block, 2015). In the negative network we model the tendency towards disliking those who are two conflicted relationships away. This effect was newly included into our version of RSiena for this article and aims at replacing a number of possible specifications that are all related to the case of emergence of negative triads.

We further control for the emergence of two-paths in both networks (connecting to those with a high outdegree). In case of the negative network this relates to the tendency to dislike others who dislike many. The next group of effects models out-degree dispersion and thus the tendency of few individuals to have many more positive or negative outgoing ties than others. We further include two effects that relate to the tendency of having zero ties rather than one or more (out-isolation) and the tendency to connect to those who have an in-degree of zero (in-isolation). In the first case we argue that in negative networks the out-isolation can be understood as not participating in the dislike network by nominating no one. The second case can, for example, relate to not disliking those who are not disliked otherwise, or to the tendency not to be someone’s only friend. The final two effects model whether individuals with a specific attribute are more likely to have more ties (ego variable), or to be nominated in the positive or negative network (alter variable).

Figs. 6–8 summarize all effects used in the SAOMs in this paper and present positive endogenous mechanisms, negative endogenous mechanisms, and co-evolving mechanism separately.

Appendix D. Discussion of newly developed RSiena effects

A new class of RSiena effects was developed for this paper. These are the configuration c in Fig. 6 and configurations e, f, g in Fig. 7. The innovation is that we can test the effect of triadic structures that involve a “conflicted tie” that is a dyad B–C in which either B dislikes C, or C dislikes
B, or they mutually dislike each other. Consider, for example, the mechanism in Fig. 8d. It expresses the tendency of individual A to dislike individual B when A’s friend C is in conflict with B. The new effect considers the three cases of i) B disliking C, ii) C disliking B, and iii) mutual dislike between B and C. These cases might be substantially distinct but they all can be argued to relate to individuals’ need to be embedded in cognitively balanced, triadic structures. The new effects are straightforward to interpret and and increase the statistical power of the model as less parameters need to be included (Stadtfeld et al., 2018).

Technically, the extension was achieved by modifying the RSienna software with a new C++ iterator for cases of dyads in which either of two ties exists. A similar iterator was added for reciprocal connections. The effect in Fig. 8d, for example, includes such a reciprocal connection. The new iterators are readily available for the development of new effects.
Stadtfeld, Christoph, 2018. The micro-macro link in social networks. Emerging Trends in the Social and Behavioral Sciences.

Stadtfeld, Christoph, Pentland, A., 2015. Partnership Ties Shape Friendship Networks: A Dynamic Social Network Study. Social Forces 94, 453–477.

Stadtfeld, Christoph, Tom A. B. Snijders, Christian Steglich, and Marijke A. J. van Duijn 2018. Statistical Power in Longitudinal Network Studies. Sociological Methods & Research.

Stark, Tobias H., Flache, Andreas. 2012. The double edge of common interest: Ethnic segregation as an unintended byproduct of opinion homophily. Sociology of Education 85, 179–199.

Stark, Tobias H., Flache, Andreas, Veenstra, René, 2013. Generalization of positive and negative attitudes toward individuals to outgroup attitudes. Personality and Social Psychology Bulletin 39, 608–622.

Tajfel, Henri, Turner, John C., 1979. An integrative theory of intergroup conflict. The social psychology of intergroup relations 33, 74.

Turner, John C., Hogg, Michael A., Oakes, Penelope J., Reicher, Stephen D., Wetherell, Margaret S., 1987. Rediscovering the social group: A self-categorization theory. Basil Blackwell.

Vörös, András, Snijders, Tom A.B., 2017. Cluster analysis of multiplex networks: Defining composite network measures. Social Networks 49, 93–112.

Watts, Duncan J., Strogatz, Steven H., 1998. Collective dynamics of ‘small-world’ networks. Nature 393, 440–442.

Yap, Janice, Harrigan, Nicholas, 2015. Why does everybody hate me? Balance, status, and homophily: The triumvirate of signed tie formation. Social Networks 40, 103–122.

Zeggelink, Evelien P.H., Stokman, Frans N., Van De Bunt, Gerhard G., 1996. The emergence of groups in the evolution of friendship networks. Journal of Mathematical Sociology 21, 29–55.