THE ROLE OF ENDOGENOUS NETWORK MECHANISMS FOR ETHNIC FRIENDSHIP SEGREGATION IN CLASSROOMS

Robert L. Hellpap
Nuffield College
University of Oxford

A thesis submitted for the degree of Doctor of Philosophy

Trinity Term 2019

Abstract

This thesis studies the interrelation of ethnic homophily and three endogenous network processes as mechanisms related to ethnic segregation of friendship networks in classrooms. Stochastic actor oriented models are applied to a German and a Swedish sample of several classrooms to investigate in-group tendencies in transitivity, reciprocity and centrality. Findings from a hierarchical analysis of the dynamics in 33 classrooms (HSAOM) from the German sample reveal that dyadic homophily can only be found for non-native pupils. Accounting for homophily in transitivity reveals natives having a small tendency for embedded ties to the out-group, while non-natives show no differences of transitivity to in- or out-group. The data suggest that homophily for non-natives might be driven by the dissolution of one-sided not-embedded friendship nominations after the first observation. HSAOM are used to analyse the Swedish sample of 194 classrooms in order analyse the role of reciprocity for the creation and maintenance of pupils' in- and out-group friendships. The results show similar tendencies for one sided and reciprocal homophily in friendship creation. While for friendship maintenance homophily is apparent in asymmetric relations, reciprocated friendships are more often dissolved within groups than between. Lastly, stochastic actor oriented models applied to separately analyse a subset of 64 classrooms from the Swedish CILS4EU data and the results are subsequently used for data-driven simulations in order to investigate how structural homophily mechanisms can have lead to the segregation of the first network observations. Results from the meta-analysis of the classroom models indicate no clear tendencies of homophily in either reciprocity, transitivity or centrality. However, accounting for these mechanism in data-driven agent based simulations provides give some insights about the emergence of segregation. The presented research projects underline that endogenous network mechanisms represent different social situations for pupils' choices about friendships to their ethnic in- and out-group. Accounting for these, helps to understand which relational structures between pupils from different backgrounds could be encouraged to avoid or reduce segregation.

THE ROLE OF ENDOGENOUS NETWORK MECHANISMS FOR ETHNIC FRIENDSHIP SEGREGATION IN CLASSROOMS



Robert L. Hellpap
Nuffield College
University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity Term 2019

Acknowledgments

First of all, I want to thank my family, no matter how many degrees of separation, for being united in encouraging me all the way from starting a career as a gardener to pursuing a doctoral degree in social sciences. Next, I feel very grateful for being able to thank my supervisor Janne Jonsson, for without his support, analytical insights and stimulating exchanges I would never have been able to start, nor ever come to an end with this project. I want to mention special thanks to my former undergrad statistics lecturer Alexander Gattig who set me on the path to pursue an academic career and who became mentor and friend. Moreover, I want to thank Enis Bicer for a spontaneous visit that lead to an extensive research idea of which parts could be realised in this thesis. During the course of working on this thesis I can consider myself lucky to have received help from a lot of excellent researchers and colleagues of which I am glad to count some as my close friends (in alphabetic order): Filip Agneessens, Per Block, Zsofia Boda, Michal Bojanowski, James Hollway, Johan Koskinen, Isabel Raabe, Tom Snijders, Christian Steglich, Christoph Stadtfeld and Laurin Weissinger. Further thanks to everyone that I had the chance to learn from in discussions and/or through osmosis during several stimulating encounters at the Sunbelt conferences, the RSiena winter schools, the Nuffield college workshop on network data collection, the CILS conferences, the Duisterbelt and of course the TUSCENY meetings. Furthermore, I want to thank Nuffield College for funding my studies by providing me with a full scholarship, delicious food and the opportunity to meet a lot of great people. And if it weren't for my wonderful girlfriend, who completes my life, I still would not know where my “home” is.

Abstract

This thesis studies the interrelation of ethnic homophily and three endogenous network processes as mechanisms related to ethnic segregation of friendship networks in classrooms. Stochastic actor oriented models are applied to a German and a Swedish sample of several classrooms to investigate in-group tendencies in transitivity, reciprocity and centrality. Findings from a hierarchical analysis of the dynamics in 33 classrooms (HSAOM) from the German sample reveal that dyadic homophily can only be found for non-native pupils. Accounting for homophily in transitivity reveals natives having a small tendency for embedded ties to the out-group, while non-natives show no differences of transitivity to in- or out-group. The data suggest that homophily for non-natives might be driven by the dissolution of one-sided not-embedded friendship nominations after the first observation. HSAOM are used to analyse the Swedish sample of 194 classrooms in order analyse the role of reciprocity for the creation and maintenance of pupils' in- and out-group friendships. The results show similar tendencies for one sided and reciprocal homophily in friendship creation. While for friendship maintenance homophily is apparent in asymmetric relations, reciprocated friendships are more often dissolved within groups than between. Lastly, stochastic actor oriented models applied to separately analyse a subset of 64 classrooms from the Swedish CILS4EU data and the results are subsequently used for data-driven simulations in order to investigate how structural homophily mechanisms can have lead to the segregation of the first network observations. Results from the meta-analysis of the classroom models indicate no clear tendencies of homophily in either reciprocity, transitivity or centrality. However, accounting for these mechanism in data-driven agent based simulations provides give some insights about the emergence of segregation. The presented research projects underline that endogenous network mechanisms represent different social situations for pupils' choices about friendships to their ethnic in- and out-group. Accounting for these, helps to understand which relational structures between pupils from different backgrounds could be encouraged to avoid or reduce segregation.

Table of Contents

List of figures and Tables....vii

1 Introduction....1

1.1 Relevance and Contribution....1
1.2 Theoretical background....4
1.2.1 Ethnic homophily and segregation of friendship networks in classrooms....4
1.2.2 Endogenous network mechanisms....10
1.3 Data....17
1.3.1 Bremen Longitudinal Study....17
1.3.2 CILS4EU....19
1.4 Statistical Method....23
1.4.1 SAOM....23
1.5 Literature....28

2 Choice and structural homophily for native and non-native pupils in classroom friendship network....33

2.1 Introduction....33
2.2 Theoretical background....35
2.2.1 Node-based mechanisms of segregation....35
2.2.2 Edge-based mechanisms of ethnic segregation....40
2.3 Data and method....47
2.4 Results....50
2.4.1 Descriptives....50
2.4.2 Hierarchical Stochastic Actor Orientented Models....58
2.5 Conclusions....63
2.6 Literature....70

3 Who becomes and who stays friends? Relating ethnic homophily and reciprocity in friendship creation and maintenance in classrooms 78

3.1 Introduction....78
3.2 Theory....81
3.2.1 Theoretical outline: segregation, homophily and network mechanisms....81
3.2.2 Homophily of tie creation and maintenance....82
3.2.3 Reciprocity and homophily....84
3.3 Data and Variables....86
3.3.1 Variable definitions....88
3.3.2 Descriptives....89
3.4 Analytical approach and Method....91
3.5 Results....94
3.6 Conclusion and discussion....101
3.7 Literature....106

4 THE structures we live in keep us with our own? In-group tendencies for endogenous network processes and their role for segregation....111

4.1 Introduction....111
4.2 Theory....113
4.2.1 Segregation and homophily – macro-state and micro-process....113
4.2.2 Reciprocity....117
4.2.3 Transitivity....118

4.2.4 In-degree centrality....119
4.3 Data and Method....121
4.4 Results....124
4.5 Discussion and Conclusions....137
4.6 Literature....141
5 Epiloge....144
5.1 Contribution and implications....144
5.2 Discussion and conclusions....148
6 Appendix A....152
7 Appendix B....153
8 Appendix C....154
8.1 Function to generate a list of siena datasets from surveydata....154
8.2 Functions to select networks suitable for single classroom SAOMs....158
8.3 Function to manipulate existing siena objects....161
9 Appendix D....171
9.1 Function to change dataformat....171
9.2 Function to calculate triad cencus by group-composition....172
10 Appendix E – Statements of co-authorship....175

List of figures and Tables

Figure 1.1 Process of network segregation.....16

Figure 1.2: Ethnic composition of classrooms by size of immigrant group..... ...22

Figure 2.1 : Composition of transitive triplets.....43

Table 2.1: Basic descriptives.....52

Figure 2.2: Measures of segregation at each wave..... ...53

Figure 2.3: Within- and between-group densities by groups and waves.....55

Table 2.2: Descriptives of intra- and inter-group dyads, twopaths and triplets.....57

Table 2.3: Linear combinations of effects from hierarchichal SAOM.....60

Table 3.1: Basic descriptives.....90

Figure 3.1: Measures of segregation at both waves.....95

Figure 3.2: Change of segregation in classrooms between waves.....96

Figure 3.3. Homophily of friendship creation and maintenance by reciprocity.....99

Table 3.2: Results of Hierarchical SAOM.....101

Table 4.1: Basic descriptives.....127

Table 4.2: Results of meta-analysis of SAOM in 64 classrooms.....128

Table 4.3: Results of meta-analysis of SAOM in 64 classrooms.....129

Figure 4.1: Average bias difference by classroom size.....131

Figure 4.2: Average bias difference by proportion of immigrants.....132

Figure 4.3: Average bias difference by observed segregation.....133

Figure 4.4: Segregation of agent-based simulation M1-M3.....134

Figure 4.5: Segregation of agent-based simulation M3a-M3c.....136

Table A1: HSAOM full model specification.....152

Table B1: HSAOM full model specification.....153

E.1 - Co-authorship declaration: Dr. Enis Bicer.….....175

E.2 - Co-authorship declaration: Dr. Isabell Raabe.....176

E.3 - Co-authorship declaration: Prof. Jan O. Jonsson.....177

1 INTRODUCTION

1.1 Relevance and Contribution

The topic of immigration has received increasing attention, not only in the public discourse and political debates but also in the research community. One of the questions that arises is the integration of immigrants into their host societies. Many ethnic groups immigrating to Europe hold lower socioeconomic status compared to natives and these disparities are often reproduced in subsequent generations (e.g. Crul & Vermeulen 2003). It can be argued that one of the core institutions able to counteract this trend is the school. While other social *foci* (Feld 1981) such as sport clubs or religious schooling can be important social spaces for youths to meet peers and form friendships, schools can provide children of immigrants with important resources that are beneficial during their life-course like language abilities and credentials. Moreover, since these institutions can be regulated by the state and pupils have to attend classes, schools promise to be of core interest for policies that aim at a successful integration of immigrant youth. However, next to research that has focussed on educational attainment of immigrant youth (e.g., Heath and Brinbaum 2014; Heath et al. 2008), investigating social integration with regard to pupils inter-ethnic friendships receives increasing attention (e.g. Jugert et al. 2018; Smith et al. 2016; Leszczensky and Pink 2015; Boda and Néray 2015; Smith et al. 2014). While inter-ethnic contact can reduce prejudice and improve out-group attitudes (e.g. Munniksma et al., 2013; Munniksma et al., 2015; Stark et al., 2013; Turner et al. 2007), the absence of intergroup relations in a society consisting of different groups can cause existing prejudice to foster (Allport, 1954; Moody,2001; Stark et al., 2013) and further restrict chances of intergroup contact (e.g. Stark 2015).

Studies looking at inter-ethnic relations in friendship networks of young adolescents

commonly find that these are less frequent than friendships between pupils of similar origin. This is often related to homophily, understood as individuals' preference to befriend similar others (Moody 2001; Baerveldt et al., 2004; Hallinan and Williams, 1989; McPherson et al, 2001; Smith et al., 2014; Verbrugge, 1977; Vermeij et al., 2009). This suggests that observing friendship networks being segregated by ethnicity results from individuals preferring to be friends with others of similar background (when controlling for opportunities and alternative individual characteristics for homophily). However, the paradigm of social networks brings with it new challenges and pitfalls in which this implicit conclusion might not necessarily hold. This is partly because data on social networks differs conceptually from classical individual data due to the interdepence of observations: an existing relation between two individuals can be the result of other existing relations in the network. The work of Grund (2014) illustrates this problem for the research on homophily and network segregation by showing that low levels of homophily during the emergence of a network can result in strongly segregated networks with the presence of transitivity. This endogenous network effect describes the well known tendency that “friends of friends become friends” and is an established mechanism often referred to as triadic closure. It is however, only one example of several different endogenous network processes that might influence emerging networks to become segregated.

Thus, it helps to clearly differentiate segregation as a macrolevel network state that can be observed at any given timepoint of a network, from homophily as a micro-level process of individuals having a preference to create or maintain relations to similar others. The phenomenon of social segregation, individuals of similar characteristics being associated with one another, has been found to occur frequently and in different social contexts. While the most prominent example stems from Schelling's model of

residential segregation (1978), applications can also been found in the study of social networks (e.g. Henry et al. 2011). One of the most striking implication of these models is illustrated by the work of van de Rijt et al. (2009) who show that systems can segregate even if individuals have a preference for diversity depending on their sensitivity to composition change in their neighbourhood.

With the increasing availability of longitudinal network data and statistical methodology to analyse them, more studies investigate homophily as a mechanism of network evolution (e.g. Jugert et al. 2018; Boda and Néray 2015). So far however, no study attempted to explicitly link the micro-mechanism of homophily with the macro-outcome of segregation while investigating the role of homophily in endogenous network processes. After all, existing network structures constitute different social situations for the individuals embedded in them. It might well be argued that these are interlinked with individuals' in- or out-group friendship preferences. For example, some pupils might be more inclined to reciprocate a friendship to one of their own, or to befriend someone from the out-group if interlinked through a mutual friend, because the existing ties of others make these friendships more attractive.

Therefore, the research question of this thesis reads: “What role do endogenous network processes play for pupils intra- and inter-ethnic friendship choices and how are these linked to networks becoming more or less segregated?”

This introductory chapter illustrates the overarching research approach of the three studies presented. It will briefly outline the theoretical approach on which all research papers are founded, describe the two datasets used, and outline stochastic actor oriented models as method used for the statistical analysis. While the papers focus on different endogenous mechanisms, they contribute different aspects to answer the research question.

1.2 Theoretical background

This brief outline of the theoretical framework underlying the presented research papers is meant as a guideline for readers rather than presenting a detailed theory. This is done in order to avoid too much repetition of lengthy theoretical reasoning about similar arguments made throughout the papers. All papers fit into the general approach as outlined below, but since they focus on different network mechanisms are limited to arguments about these. Thus, this section serves as means to indicate how the different approaches share a substantive theoretical foundation.

1.2.1 Ethnic homophily and segregation of friendship networks in classrooms

The social relations that tie individuals to one another can be manifold in their nature and the consequences they have on their lives. Close relations like friendships can provide resources, support and advice (e.g. see Furman and Buhrmester, 1992; Hartup and Stevens, 1997; Helsen et al., 2000; Uehara, 1990) and satisfy needs for social acceptance and belonging (e.g. Baumeister and Leary, 1995; Denissen et al., 2008). For members from minority groups that have a disadvantageous socio-economic position in a society, such as immigrants, friendships to members from the native majority can be especially beneficial for their integration into the host society.

For children with an immigrant background, classrooms are an important context providing opportunity for intergroup friendships to form. However, research on adolescents friendships finds consistent patterns of pupils rather being friends with in-group members (e.g. of the same ethnicity) than with those from the out-group (e.g. Smith et al. 2014; Smith et al. 2016). The work presented here contributes by improving the understanding of mechanisms on the micro-level of pupils that take place in the emergence of ethnic segregation on the macro-level of classroom friendship networks. One of the core mechanisms discussed in this work refers to *homophily*, individuals'

tendency to befriend other in-group rather than out-group members (e.g. McPherson et al. 2001).

Two individuals naming each other as friends is understood as the outcome of a sequence of successful interactions they have with each other. Interactions are understood as successful, in line with social exchange theory (Emerson 1976; Rusbult and Buunk 1993), both individuals invest resources and perceive their own investments as beneficial. Before initiating an investment, individuals hold existing expectations about the certainty with which their investment into the *alter* will pay off (*interpersonal uncertainties*). The less certain a person is about the positive outcome of their investment, the higher the subjective costs of investing into the respective option of forming a relationship. This implies that while individuals make decisions about whom to interact with, they have expectations about how the others decide themselves. Thus, one would not invest into a social bond if one thinks that the other will not be interested in investing themselves. During the course of subsequent interactions, the associated uncertainties are updated and the respective costs re-evaluated. With repeating successful interactions, the subjective quality of the relationship improves until it reaches an individual threshold after which a person would categorize it as *friendship*. This approach provides a flexible framework that does not rely on individuals having an accurate network cognition, i.e. who a pupil nominates as a friend does not depend on other's friendship nominations *per se*, but rather on their perception of these. A pupil nominating another as their friend is not necessarily because the other nominates them in return. It is sufficient that a pupil might think that the other thinks of them as a friend. Thus, a friendship nomination can be understood as an indicator for a latent factor that is the result of repeated interactions, varying by the perception of both individuals that are part of these. The longer a relationship between individuals persists and the more

often they interact, the more congruent their understanding of the relationship should become.

Considering time, it becomes clear that the mechanisms for becoming and staying friends are of different nature. During the initial or early interactions between two individuals one would expect that more salient attributes play a stronger role since individuals don't know each other yet. Pupils' decisions about with whom to spend their time are more likely to be influenced by the others' gender, migration background or demeanour. With continuing interactions, individuals learn more about each other, adapt expectations, re-negotiate roles and re-evaluate the relationship. During the course of this process, one would expect that the relevance of characteristics shifts to those dimensions that are less salient. While initially two pupils decide to spend time with each other because they have the same gender, ethnicity and have the same type of humour, they later might discover that they are fans of two opposing football clubs, or have very different taste of music. Thus, over the course of time, initially hidden differences between two individuals might be revealed which put additionally strain or costs on the relationship. If at one point, a person becomes uncertain about their investment into the friendship, or perceives the associated costs as too high, their perceived relationship quality reduces to a point where they would stop to classify it as a *friendship* or even develop a *negative* relation.

In this work, ethnic homophily is understood as a systematic difference in individuals rather becoming and staying friends with in-group rather than out-group members, and heterophily for the opposite respectively. Based on social identity theory (Taijfel and Turner 1979), individuals associate themselves with social groups as part of their identity formation, and ethnicity serves as an important characteristic for group identification. Individuals tend to often associate ethnic in-group members with more

positive attitudes while out-group members are often ascribed with less positive or even negative attitudes and prejudice (e.g. Bettencourt et al., 2001; Pinter and Greenwald, 2011; Scheepers et al., 2006; Tajfel, 1970). This can lead to additional uncertainty about investments into a relationship with an *alter* from an out-group *(intergroup uncertainties*), thus intergroup relations become high cost situations compared to intragroup relations. Established friendships between members of different ethnic groups can also face higher associated costs than relationships between members of the same group. This could be due to third party influence such as other in-group peers (e.g. Kreager, 2008) or parents that discourage or sanction intergroup relations (e.g. Munniksma et al., 2012). Hence, *homophily* is understood as individuals showing a tendency to establish and maintain friendships with in-group rather than out-group members because of lower associated costs with the first compared to the latter. This can apply to different types of salient grouping characteristics (e.g. gender, age/seniority), but here the focus is on ethnic groups. In this work, factors that influence this discrepancy in costs with regard to altering individuals preference for in-group friendships are here referred to as *group boundaries*. This can include existing personal attitudes and prejudice about the other group, identification with one's own group, third party influence or even open conflict between different ethnic groups in the classroom.

Accounting for the time perspective of ongoing personal relationships, the salient characteristic of migration background could be expected to mainly be important for initial or early interactions, and should loose relevance once differences are overcome. Following, one would expect that inter-ethnic friendships are less frequently created, but similarly maintained when compared to intra-ethnic friendships. Still, it might be that differences in ethnicity come along with other less salient differences that are revealed over time, such as religious or cultural values and practices that put additional

strain on inter-group friendships. For example, an adolescent pupil that believes in Islam might be put off by their native friend trying to encourage him to drink alcohol.

School classrooms pose an ideal context for the analysis of homophily on the micro-level, and segregation on the macro-level. Pupils spend a large proportion of their time in school and thus with all pupils in their classroom. From the perspective of Feld's *social foci* approach (1981) classrooms can be understood as a restrictive setting that focusses individuals' activities with each other. Individuals that share social foci are more likely to have interactions that enable them to create positive associations with alters as a basis for establishing friendships, Pupils in a classroom are aware of, and interact with, each other on a daily basis. For the investigation of *homophily* as a process of individuals choosing with whom they rather become and stay friends with within the outlined theoretical framework, this has the advantage that pupils know all options of alters they could become friends with. However, one has to acknowledge that pupils might share other foci outside of the classroom or school context. For example, they could be spending time with each other in the same sport clubs, musical groups or religious classes. Moreover, these foci can be segregated based on group membership, because groups might have different preferences for participation in specific foci (Feld 1981). This can lead to the problem that apparent ethnic homophily or segregation in a classroom might be the result of pupils' from the group of natives or non-natives sharing *foci* within the respective group. Thus, focussing only on the integration of immigrants within the school context can potentially be misleading about possible causes of segregation or homophily. One can argue that ethnic friendship segregation in classrooms mainly expresses that *foci* are rather homogeneously composed of adolescents with similar migration background. In those alternative *foci* pupils mainly share activities with their own and thus rather befriend each other. This being the case

might result in wrongly attributing causes of segregation, but would imply that the segregation or integration within a classrooms is a proxy for the friendship segregation outside the school context. For example, it might be that two ethnically distinct friendship clusters form in a classroom, not because of pupils' having a high preference for ingroup friends, but because of pupils in that classroom are members of two different football clubs. If neighborhoods are ethnically segregated, natives and non-natives respectively play in the club of their neighborhood, the local clubs are likely segregated too. Thus, the ethnic segregation one would observe in the classroom reflects the segregation of two overlapping *foci*, neighborhoods and football clubs.

Analysing *segregation* of friendship networks in school classrooms is straightforward as one can collect information about complete networks within these boundaries. *Ethnic* S*egregation* is here understood as a network state with ties between actors existing predominantly within ethnic groups, than between them, and *integration* as the opposite case[1]. In this work, this differentiates it clearly from the concept of *homophily* which relates to the mechanisms on the micro-level leading to the observable network state *segregation*. Additionally, setting the network boundaries at the classroom level also allows to control for the *opportunity structure*. It is one of the key determinants of inter-ethnic friendships to form at all because it defines the availability of ethnic out-group members to possibly befriend. Moreover, for school classrooms the information about the composition of pupils based on their migration background is available, which allows to investigate segregation and homophily net of *opportunity*. Thus, both concepts relate to in- and out-group friendships accounting for the possible options of friends they have within the classroom. Additionally to determining the opportunities for native and non-native pupils to interact and potentially become friends, the ethnic composition

1 This does not mean that segregation measures can only be calculated for the network level (compare Bojanowski and Corten, 2014).

of classrooms might moderate pupils tendency to befriend ingroup alters (e.g. Smith et al. 2016). Ethnic homophily depends on the salience of pupils' migration background, because the absence of apparent grouping characteristics would be expected to make friendship choices independent of migration background. It can be argued that the more heterogeneous the ethnic composition in a class the less salient grouping characteristics are perceived. If all pupils have a different migration background, their respective grouping characteristics should become less relevant and other dimensions for identification and distinction would be expected to be more important for friendship formation.

1.2.2 Endogenous network mechanisms

Ethnic homophily, understood as a micro-level mechanism during the process of the evolution of a dynamic social network, can lead to ethnically segregated networks. However, observation of a strongly segregated network does not necessarily imply that actors in the network had high levels of homophily. This is (at least partly) because of the interdepence of observing social ties in a complete network. Actors – and the ties connecting them to others – are also embedded within the surrounding network of social relations between alters. An isolated pupil that becomes friend with another who already has a lot of friends, gets indirectly connected to several new others. When investigating the evolution of social networks, several different types of *endogenous network mechanisms* have been identified that can differ strongly between type of the network, actors, or context. In a dynamic perspective, these mechanisms take place such that specific network structures have a tendency of transitioning into specific other structures (e.g. Balance theory, Davis 1963).

Reciprocity is one of the most commonly established network mechanisms for friendship networks. Relating to social exchange theory (Emerson 1976; Rusbult and

Buunk 1993), reciprocated friendships can be considered as being perceived of higher quality compared to those friendships into which only one pupil invests. Thus, during the process of repeated interactions, if individuals felt that their investment into the relation is not reciprocated, they might decide to cease investing before an actual friendship is established. In this case, one would expect that most friendships emerge reciprocally or don't emerge at all. However, one could observe one-sided friendship nominations if both actors in a dyad have different thresholds of categorizing the social relation as friendship. Thus, both pupils might invest to an extent that they perceive as beneficial, but one of the pupils considers their relation as a friendship sooner than the other. Established reciprocal friendships would also be expected to exist longer since they are more beneficial for both individuals compared to asymmetric friendships. With regard to homophily, it is argued that in-group dyads pose a lower cost scenario than out-group dyads. Thus, the latter would need to compensate involved actors by providing additional benefits, or perceived gains such as reciprocation. The other way around, one could argue that friendships within groups do not necessarily need to be reciprocated, since the shared group identification accounts for such an imbalance. This would imply that asymmetric friendships rather emerge within groups than between, while reciprocal relations rather emerge between groups. For the emergence of segregation, this can mean that friendships between groups are initially less likely to emerge, because they are more prone to be dissolved if perceived differently, or asymmetric by both actors. Thus, networks initially segregate because of few reciprocal ties being established between groups, while asymmetric relations are formed within.

In contrast, focus theory (Feld 1981) suggests that individuals who share *foci* mutually perceive relationships as positive or beneficial. Thus, if homophily results from shared *foci* within groups, one would expect to find more reciprocation between pupils of the

same migration background.[2] In that regard, reciprocated friendships would rather emerge within groups, and as a result of fewer shared *foci* between groups one-sided friendships are more likely to emerge between pupils from different migration background. This means, if actors show homophily of reciprocity, i.e. they establish more reciprocal ties to the in-group, this would lead to a much stronger rise of segregation than in the first case.

Regarding the maintenance of friendships, reciprocal friendships are more stable or less prone to being terminated (e.g. Gershman and Hayes 1983). If this tendency is similar within and between groups, this means that classes which are ethnically segregated because of shared social *foci* within groups would be expected to stabilise at higher levels of segregation or even increase. In the opposite case, initial segregation could be expected to reduce because once mutual friendships have formed between groups they would be expected to persist, while asymmetric in-group relations are more prone to fluctuation.

The second endogenous network mechanism under investigation here is *transitivity*, individuals becoming friends with the friends of their friends. The tendency for such triadic network structures is a common finding in adolescents friendship networks. Having a mutual friend brings other alters closer in the personal social space. With an existing good friend, ego already overcame some initial *inter-personal uncertainties* during the friendship formation process and have established a certain level of trust towards the respective alter. The same holds for the friends that alters make, thus suggesting that an investment into a relationship with them is more likely to succeed, because the already existing friend indirectly vouches for them. Given that time is a limited resource, and an ego wants to spend leisure time with an alter, it is likely they

2 More precisely, Feld (1981) explicates that two individuals (mutually) connected to each other share *foci*. If they again are connected to third actor, it is likely that those share the same *foci* as well. Thus, he follows that that mutually connected triplets are the result of more individuals sharing the *foci*.

will also spend time with the alters' friends. As has been pointed out by Grund (2014), the tendency for transitive closure can amplify low levels of homophily in order for networks to become strongly segregated (based on network size).

Following the arguments of Block (2015), establishing and maintaining relations within groups pose lower costs for individuals as is the case between groups.

groups of three (triads) can be seen as a small group and thus pose lower cost situations compared to non-embedded dyadic links. On the one hand, this could imply that triadic relations counteract ethnic boundaries to some extent, thus leading to more inter-ethnic ties to be created through interconnected pairs. It can even imply *heterophily* in transitivity, meaning that indirect friends are more likely to become direct friends when they are from the out-group (*inter-group transitivity*) rather than being from the same group (intra-group transitivity). This is because for other pupils from the own group, having a shared friend is not necessary to become friends because of identification with the same ethnic group. On the other hand, with the presence of strong group boundaries where individuals put a strong emphasis on identification with their own ethnic group, intra-group transitivity might provide an ideal means to enforce group distinctions. Thus, in order to receive additional personal validation, individuals might strive for ties to their own group that are embedded in triplets resulting in proportionally more embedded in-group ties being realised than embedded ties to the out-group (*transitive homophily*)

Following the arguments of social focus theory (Feld 1981) suggests that individuals from the same migration background are more prone to sharing multiple *foci* and as a result lead to (mutually) connected triplets. Following, one would expect to find homophily in transitivity, referring to two-paths being more likely to close between actors of the same migration background.

For the first case, this could mean that segregation rises in the presence of (some) dyadic homophily and a tendency for transitivity, but might be moderated to some extent due to more two-paths closing between pupils of different background. In the opposite, social *foci* proclaim that pupils show a preferences for closing triads rather within groups, which would lead to an even stronger increase of segregation. If one argues that transitive triplets are more stable, due to an enforcing or confirming role of the intermediate actor, this would again imply more stable or increasing levels of segregation for the case of social *foci* and a reduction of segregation otherwise.

Lastly, this work investigates how *centrality* relates to homophily and segregation. The concept of centrality has long been established in research on different types of social networks, for example flow networks (Borgatti 2005), affiliations (e.g. Faust 1997) or gossip and friendships (Ellwardt et al. 2012). In the context of this work, a pupil is understood as being more *central* when receiving more friendship nominations than most others in the network. Pupils that are popular in this manner, can be seen as facilitators since they interconnect several alters and provide opportunity for transitive closure. However, there is a limit to the amount of close friendships one can sustain (Dunbar 2004) such that central actors might not be able to invest into all of the relations to alters who consider them as friends. This can lead to pupils centrality stopping to increase, or even decrease at a certain point since alters consider their investments into that relationship as not beneficial for themselves over the course of time. But, pupils are assumed to associate lower costs with in-group compared to out-group relationships, and would rather continue to hold not reciprocated relations to the first. From this it follows, that pupils would be more likely to become central within their own groups since incoming friendships from the out-group would be dissolved earlier if not reciprocated. Pupils showing to have a tendency of rather creating and

maintaining friendships to central actors of their own group display *homophily of centrality*. This would imply that individuals showing to be connected to similar others, might not be the result of an ingroup preference for friendships, but rather because they are connected to similar others who are central within their own group. Since, central actors can be understood as facilitators for transitive closure, actors that are predominantly central within their own ethnic group contribute to an increase of segregation because they indirectly connect several pupils from the same migration background.

However, taking social *foci* into account one would not necessarily expect homophily in centrality, because it predicts mutually transitive ingroup clusters. With that respect, actors could be more central between groups because they might share some *foci* with outgroup members, though less than they do with their own. Thus, actors are central because they partake interactions in some *foci* that are predominantly constituted by their outgroup.

If actors become more central within their own group, this would be associated with increasing segregation, but if they are similarly central between groups they could counteract other segregating mechanisms.

One possible extension of the arguments about the social situations that network structures provide, goes beyond the context of the classroom. Following the approach of Leszczensky and Pink (2015) one can argue that friendships between pupils in different classrooms are faced with higher cost than within classrooms, because the latter are a more restrictive *foci*. They follow that homophily should be more pronounced between classrooms at the grade level, and find evidence supporting their expectation. Similarly, this argument can be made for the three introduced network mechanisms. Generally, one would expect that that they are more pronounced outside the classroom. For example,

transitivity might be stronger because interconnected alters provide a sort of pre-selection of alters to connect to. Thus, it might be possible that the finding of stronger homophily between classrooms is to some extent mitigated by strong transitivity between classes. While these arguments are not further investigated within this work, it is worthwhile considering their implications for the validity of findings presented here.

The theoretical framework can be summarised in line with analytical Sociology (Hedström 2005) by projecting it in a *Coleman Boat* (Coleman 1990). Figure 1.1 illustrates how the macro-level characteristic of segregation in a classroom friendship network can emerge and change. Imagining a network at t0 in which no ties between actors have been created yet (top left in figure 1.1), which is found to be segregated at t2 (top-right). The composition of this network – with regard to the number of actors belonging to one group or the other – constitutes the opportunity structure (arrow 1) for pupils' friendship choices. Based on the opportunities for each actor to befriend in- or out-group members we can calculate the expected existence of ties if actors randomly choose new friends (i.e., assuming no homophily). If proportionally more ties are created within rather than between groups this is taken as an indicator for homophily (arrow 2), assuming that actors rather choose new friends from their own group.

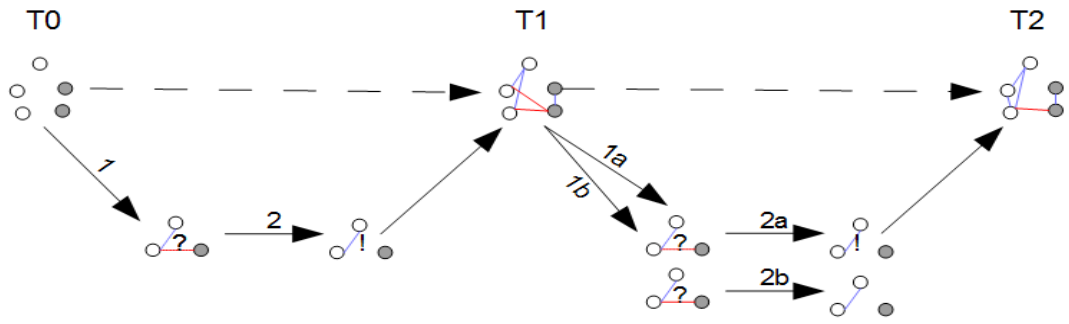


Figure 1.1 Process of network segregation

While this might be the case at the beginning of an evolving network, observing more in-group than between-group ties at later stages (T1 to T2) could also result from actors

having more indirect ties to in-group members and simply become friends of their friends without any actual in-group preference. It is therefore important in a study of segregation and homophily to take network-inherent mechanisms into account (arrow 1a), such as transitivity that occurs during the evolution of a network. Moreover, the current network structure as well as its composition can have an effect of actor's preferences for in- and out-group ties by enforcing group boundaries (arrow 1b). Importantly, for a network to become segregated further it is not necessary that more in-group ties are created (arrow 2a) – this could also be achieved by in-group ties persisting at least as long as or longer than between-group ties (arrow 2b). In order to better understand the processes leading to segregated (friendship) networks, it is important to analyse the change of networks over time and investigate group-specific tendencies of both tie creation and preservation.

1.3 Data

1.3.1 Bremen Longitudinal Study

The first data stems from the DFG-funded project "Integration through friendship? Dynamics of social assimilation of children from immigrant families in multiplex peer networks"[3]. It was conducted in two German cities, namely Bremen and Bremerhaven, located in the north-west of the country with both having a considerable proportion of inhabitants of non-German immigration background. The goal was to gain insight into integration dynamics of immigrant children. One of the core interests was their social integration and its effects during the early life-course. Therefore, the study collected information on several dimensions of social relations within the classrooms, as well as outcomes of educational achievement and personal well-being. Before the start of the longitudinal dataset, one cross-sectional wave was collected in a total of 105 classrooms

3 DFG grant no. WI 3423/1-1, AOBJ 574807

taking place in 4th grade, the last year of primary schooling in Germany. The main-phase of data collection took part during three consecutive years starting in 2010 half a year after the beginning of fifth grade which is the first year of secondary school and subsequent waves taking place every following year. Participation of pupils needed three stages of opting in from heads of the school, the teacher and parents. From the 55 schools that have been approached, 35 (63.6%) agreed to participate in the first wave, in which 94 classrooms (63.1%) provided information of 1676 pupils (75.4%). However, the participation varies over the course of data collection, since some schools stopped their participation while others joined. This leads to number of participating classrooms dropping to 58 in the second wave and 65 in the third. In total this left 22 classrooms which participated in all three waves, and 37 participating in at least two waves. From the all pupils participating at least in one wave of the four data collection points (n=6606), 57.1% report that both parents are of German origin, 13.8% report one parent with immigrant background and 29.1% of the pupils report that both their parents have an immigrant background (Bicer and Windzio, 2014).

Several dimensions of network data were gained from sociometric instruments in the questionnaire. In order to ensure anonymity of the participating pupils, during the collection of the sociometric questionnaires, each pupil was assigned a unique identifier which was printed on a flag next to each participants' table. When filling out the items of the sociometric questionnaire, pupils could then enter the unique identifiers of their respective peers. In total 17 network dimensions have been collected for the three waves. This includes positive dimensions such as friendships, birthday invitations, helping each other with homework or spending time on the schoolyard. Negative relations like having a dispute/problems, not wanting to sit next to, say nasty things, or having been kicked/beaten by. Lastly, some instruments collected rather informative

relations about proximity of living places or whose parents know each other.

1.3.2 CILS4EU

One major goal of the CILS4EU project (Kalter et al. 2016; 2018) was to gain insight into the process of cultural, social and structural integration of pupils with immigrant background into their host societies, and help uncover mechanisms leading to inequalities. In order to achieve this, longitudinal data of native and non-native adolescents (initially) at the age of 14 was collected. Furthermore, the study allowed comparing adolescent behavior between the (initially) four participating European countries Germany, United Kingdom, The Netherlands and Sweden. In order to have a sufficiently large sample of non-native students to allow for systematic comparisons to the reference group of native pupils, a stratified three-stage sample design was realised. The first stage selected schools with probabilities proportional to size, oversampling those with higher proportions of pupils with immigrant background. The second stage targeted two randomly sampled classes within these schools, while the sampling units of the third stage are all students within these classes. The first wave of data was mainly collected in the spring of 2011 and included student, parent and teacher questionnaires. Hereby, the parent and teacher data was collected as supplementary cross-sectional data during the first wave. The student survey included a wide spectrum of questions in order to capture relevant dimensions of structural, cultural and social integration, established and possible intervening and explanatory mechanisms, a test of cognitive and language ability and a sociometric questionnaire. This dataset provides an unprecedented opportunity to study inter-ethnic friendships and their dynamics within classrooms while accounting for different mechanisms influencing these.

The dataset provides two types of network information, one being ego-centered networks based on a name generator asking pupils about their five best friends as well

as basic characteristics of these, like age, gender, ethnic migration background, education, meeting places, and if their parents know them. Pupils could name friends outside of the school as well as others from their classrooms. The second instrument was a sociometric questionnaire capturing complete network information within the boundaries of the classroom. Each student received a list showing all names of the other students and a corresponding unique identifier. Students made nominations by listing the IDs of all alters they are related to per specific network dimension. The sociometric questionnaire included a total of 11 instruments that can be categorised as positive, negative and informative relations. Positive relations asked for pupils five best friends, the single best friend, who the most popular pupils are, who they spend time with outside of school, and who they do homework with. Negative relations relate to who pupils do not want to sit by, who is mean to them, who they are mean to. Informative relations ask about which classmates live within a 5 minute walk, who their parents know and whose parents their own parents know. However, not all dimensions were asked at the second wave. Negative relations were dropped from the questionnaire entirely and the dimension asking about proximity of the place of living has not been updated either.

The work presented here will focus on pupils' nominating their five best friends. On the one hand, this has the limitation of pupils being restricted to name a maximum of five outgoing nominations, it presents the strongest positive relation. This promises to pose the largest contrast of pupils in- and out-group friendships, which is crucial since these close friendship relations are also the ones most beneficial, especially in their integrating function for non-native pupils. Also, the friendship dimension is likely to cover the others, meaning that another pupil who is your close friend will also be someone you spend time with during breaks or outside school.

The research in this work, is based only on the Swedish sample with 4108 pupils from 127 schools participating in the first and the second wave consisting of 47.2% of non-native pupils. The immigrant group in Sweden is diverse with regard to their origins. This is also reflected in the sample, the biggest group having an immigration background of Arabic, Eastern Europe, other parts of Europe, and African or Asian origin. It is plausible that classrooms with higher proportions of non-native pupils also have the potential for a more diverse composition of pupils' immigrant backgrounds. As illustrated in figure 1.2, this relation is not linear. While the diversity in classrooms increases with the amount of non-native pupils, those classrooms where immigrant children pose the majority have more variance in their composition. It follows that some of the classrooms with a non-native majority of pupils consist of immigrant children with a variety of different backgrounds, while others consist of only few different backgrounds or potentially just one ethnic group.

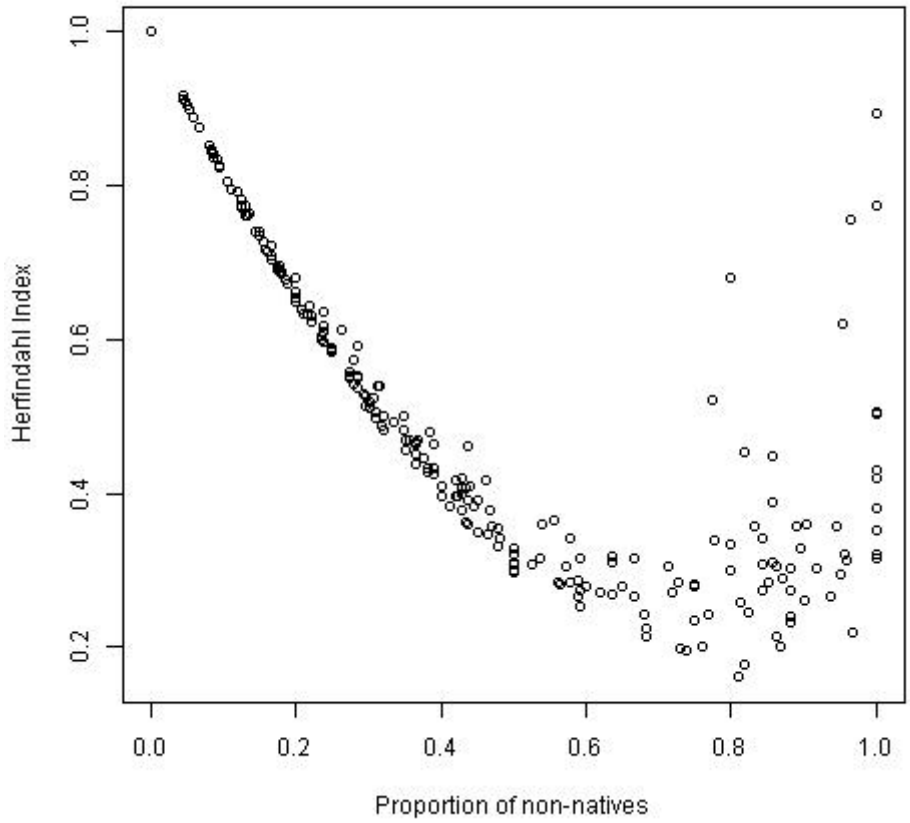


Figure 1.2: Ethnic composition of classrooms by size of immigrant group

Compared to the other countries, analysing the Swedish sample of the CILS4EU project has advantages. One is that the information provided by respondents has been cross-validated with data from the Swedish population register, improving data quality, for example for parental educational background. One major benefit relates to the sociometric questionnaire, which compared to Germany and The Netherlands does not provide so much change of the pupils constituting the classrooms. One of the reasons for this is that in the Swedish school system, pupils rarely if ever repeat classes, such that in this data only few of them leave or join into the classrooms between data-collection points. Another reason is that these two grades (8^{th} and 9^{th}) are both within the comprehensive school system, meaning that there is no possibility of moving between

tracks, as in the Dutch or German systems. The pupils persistence over time is especially important for the longitudinal network analysis since estimations will be based on more datapoints.

The research approach used in this work has not been attempted before, especially with such a big dataset. For coding the large amount of classroom networks, several functions in the R statistical package have been created and are being prepared to be made available as a package on R-forge. The goal is to make existing tools for social network analysis more accessible for researchers having access to survey data including sociometric questionnaires (see Appendix C). Since the attempted research proved to be highly complex, focusing on one country alone keeps the scope in a tangible reach and leaves country comparisons for subsequent research.

1.4 Statistical Method

1.4.1 SAOM

The baseline statistical procedure applied in this work are stochastic actor oriented models (SAOM) (Snijders et al., 2010b) to analyse changes of longitudinal networks. To receive parameter estimates, the first observation of a network is used as a starting point for the estimation, which is carried out by simulating network ties (and behaviour) in a continuous time Markov process (Steglich et al., 2010). The SAOM models the network evolution in continuous-time, assuming that actors make changes to their ties at different points in time. The main target of inference is to model these changes as choices of ego *i* in so-called *mini-steps*. Each of these mini-steps represents an individual choosing whether a new tie will be created or an existing tie will be dissolved. In the longitudinal data used in this study, the mini-steps are not directly measured, meaning that the exact sequence of decisions in-between waves is unobserved. Nonetheless, the estimation of

model parameters can be conducted applying any of three estimation procedures that account for this type of missing information: method of moments (Snijders, 2001); Bayesian (Koskinen and Snijders, 2007) and maximum likelihood (Snijders et al., 2010a).

The SOAM provide an analytical tool that allows to adequately represent the theoretical framework, since pupils have agency in making choices about their friendships as well as take into account the existing network structure and it's endogenous tendencies to change.

It should be noted that SAOMs take the first observation as a starting point, which might be problematic for this research agenda if the network mechanisms that are important for the emergence of segregation already took place. In other words, it might well be that the first observation is the result of the introduced endogenous network processes. As such, investigating systematic change towards the second observation might not result in gaining insights about the emergence of segregation, but mainly about the dynamic of it. This could be especially problematic for the case of the Swedish classrooms, since their composition of pupils can be considered as rather stable (i.e. they know each other for longer, had more opportunity to interact and therefore most network structures have already been established). One can argue that this problem might be addressed by conducting cross-sectional analysis of the first network observations, such as ERGMs (Lusher et al. 2013). Still, these approaches might be misleading, because they depend on the observed network structures which could be the result of different types of network mechanisms. For example, if the first network observations show a high proportion of ethnically homogeneous triads this could be due to different mechanisms having taken place:

1. Low levels of dyadic homophily and presence of (general) transitivity (following Grund 2014)

2. Low to no dyadic homophily, but mutually shared *foci* (following Feld 1981)

3. Low to no dyadic homophily, but homophily in centrality and the presence of transitivity.

While it might be impossible to differentiate these processes without having continuous time data that starts from observing a network at an initial (empty) starting point, one attempt is made in the last chapter of this thesis.

Parameters estimated in SAOMs express changes within a network as the result of individuals choosing to alter their outgoing ties. Since the stochastic actor oriented models can be regarded as an agent based model (Macy and Willer, 2002), applying simulations of micro-level mechanisms is straightforward (Snijders and Steglich, 2015; Prell and Lo, 2016; Block et al., 2018; Stadtfeld et al., 2018). In order to investigate possible mechanisms for the emergence of segregation in a given network, estimated sets of parameters are used to simulate the evolution of a given network. The starting points for these simulations are empty networks with the number of actors present during the first observation. The simulation then processes a sequence of mini-steps based on multiples of the average changes between observations. After each of these mini-steps, the level of segregation in the resulting network can be calculated to monitor the macro-level outcome influenced by the micro-level mechanisms. Comparing the resulting simulated networks with the observed networks can provide insights about the processes that took place during the emergence of the observed segregation. The more similar the segregation level of the simulations is to the observed segregation, the more plausible it is that the micro-level mechanisms included in the estimation played a role

during the emergence of the observed network state. If one applies this method as means to potentially explain how different mechanisms can be the cause for the first network observation, one has to be clear that this is under the assumption that the network dynamics are constant. Of course, one point made earlier is that the mechanisms for tie creation and maintenance are different by nature. For the emergence of segregation from an empty network firstly tie creation has to take place, before tie maintenance can play a role at all. Of course, this strongly challenges the assumption about stable network mechanisms. Therefore, the work presented in chapter 4, should rather be understood as a starting point that introduces the potential benefits of the applied method for the outlined problem in this context. Work that builds on this approach should firstly account for the fact that network mechanisms are different during initial emergence and later dynamics.

Having repeated observations on multiple independent networks allows to apply a hierarchical model that can account for heterogeneity between classrooms. This is possible with Bayesian Hierarchical SAOM (HSAOM), the recently developed extension of classical SAOM models (Koskinen and Snijders, 2018). The HSAOM can fit the same model specifications to all classrooms, which otherwise proves difficult in several cases, for example because of the presence of some small classrooms. It is conceptually similar to classical regression-type multilevel models with regard to differentiating variation on the individual and on a group level. In the HSAOM approach, each classroom g=1,...,G is assumed to follow the same SAOM but with a classroom-specific set of parameters $\theta^{[g]}$. These parameters in turn are assumed to be drawn from a population-level model which is a multivariate normal distribution with a mean μ and covariance σ. The target of inference is primarily to estimate the population mean μ rather than the classroom-level parameters $\theta^{[g]}$.

The estimation-procedure involves drawing samples of three key posterior variates: the population parameters, the classroom-level parameters, and the sequences of unobserved changes in-between waves. These variates are generated through an iterative Markov-chain Monte Carlo (MCMC) sampling scheme. Inference is based on these draws from the posteriors given observed data. From a non-Bayesian perspective, the posterior mean and standard deviation of parameters can be interpreted as approximations to the MLE and standard error, respectively. In the equivalent of fixed effect models, between classroom variation of the classroom-level parameters[4] can be fixed to be the same $\theta^{[g]}=\theta^{[h]}$ for all g and h[5]. All SAOM applications are carried out using the RSiena Software (Ripley et al., 2018). The work on this thesis contributed to the further development of RSiena by programming new effects that have been added to the software. Additionally, several functions (see Appendix C) have been written for preparation and manipulation of datasets with information on multiple social networks to ease the application of RSiena. These functions are in preparation to be submitted as a publicly available package on R-Forge.

4 Except for outdegree

5 While g represents the overall parameters, h represents the individual classroom parameters.

1.5 Literature

Allport, G. W. (1954). The Nature of Prejudice. Addison-Wesley Publishing.

Baerveldt, C., van Duijn, M.A.J., Vermeij, L., and D.A. van Hemert, Dianne A. (2004). Ethnic boundaries and personal choice. Assessing the influence of individual inclinations to choose intra-ethnic relationships on pupils´ networks. *Social Networks* 26. 55–74.

Baumeister, R.F. and M.R. Leary. (1995). The Need to Belong: Desire for Interpersonal Attachments as a Fundamental Human Motivation. *Psychological Bulletin* 117 (3), 497–529.

Bettencourt, A. B., Dorr, N., Charlton, K.; and D.L. Hume. (2001). Status differences and intergroup bias: A meta-analytic examination of the effects of status stability, status legitimacy and group permeability. *Psychological Bulletin.* 127. 520–42.

Bicer, E. and M. Windzio. (2014). Soziale Integration durch Kanten zwischen Knoten. Methoden und Daten in der netzwerkanalytischen Integrationsforschung. Enis Bicer, Michael Windzio und Matthias Wingens (Hg.): Soziale Netzwerke, Sozialkapital und ethnische Grenzziehungen im Schulkontext. Wiesbaden: Springer Fachmedien Wiesbaden GmbH, S. 75–99.

Block, P., J. Koskinen, J. Hollway, C. Steglich, and C. Stadtfeld. (2018). Change we can believe in. Comparing Longitudinal Network Models on Consistency, Interpretability and Predictive Power. *Social Networks.* 52. 180–91.

Boda, Z. and B. Néray. (2015). Inter-ethnic friendship and negative ties in secondary school. *Social Networks.*43. 57-72.

Bojanowski, M. and R. Corten. (2014). Measuring segregation in social networks. *Social Networks*. 39. 14-32.

Borgatti, S. P. (2005) Centrality and network flow. *Social Networks.* 27(1). 55-71.

Coleman, J. (1990). Foundations of social theory. *Harvard University Press*

Davis, J.A. (1963) Structural Balance, Mechanical Solidarity, and Interpersonal Relations. *American Journal of Sociology.* 68 (4). 444-62.

Denissen, J.J.A., Penke, L., Schmitt, D.P., and M.A.G van Aken. (2008). Self-Esteem Reactions to Social Interactions: Evidence for Sociometer Mechanisms Across Days, People, and Nations. *Journal of Personality and Social Psychology.* 95 (1). 181–96.

Dunbar, R.I.M. (2004). Gossip in evolutionary perspective. *Review of General Psychology.* 8, 100–10.

Ellwardt, L., Steglich, C. and R. Wittek (2012). The co-evolution of gossip and friendship in workplace social networks. *Social Networks.* 34(4). 623-33.

Emerson, R. M. (1976). Social exchange theory. *Annual review of sociology*, 335-62.

Faust, K. (1997). Centrality in affiliation networks. *Social Networks.*19 (2).157-91.

Feld, S. (1981). The Focused Organization of Social Ties. *American Journal of Sociology, 86*(5), 1015-35

Furman, W., and D. Buhrmester. (1992). Age and Sex Differences in Perceptions of Networks of Personal Relationships. *Child Development* 63 (1), 103–115.

Grund, T. (2014) Network Size and Homophily: Same-Sex Friendships in 595 Scandinavian Schools, in Manzo, G. (eds.) (2014) Analytical Sociology. Norms, Actions, Networks. New York: Wiley. 299-316.

Hallinan, M.T., and R.A. Williams, (1989). Interracial Friendship Choices in Secondary Schools. *American Sociological Review.* 54 (1). 67–78.

Hartup, W.W. And N. Stevens. (1997). Friendships and Adaptation in the Life Course. *Psychological Bulletin.* 121(3). 355–70.

Heath, A.F. and Y. Brinbaum. (2014). Unequal Attainments: Ethnic Educational Inequalities in Ten Western Countries. Proceedings of the British Academy, 196. Oxford: Oxford University Press.

Heath, A. F., Rothon, C., & Kilpi, E. (2008). The Second Generation in Western Europe: Education, Unemployment, and Occupational Attainment. *Annual Review of Sociology*, *34*(1), 211–235.

Hedström, P. (2005). Dissecting the social. On the principles of analytical sociology. *Cambridge: Cambridge University Press*.

Helsen, Marianne; Vollebergh, Wilma; Meeus, Wim (2000). Social Support from Parents and Friends and Emotional Problems in Adolescence. *Journal of Youth and Adolescence.* 29 (3). 319-335.

Henry A. D, Pralat, P., and C.-Q. Zhang. (2011). Emergence of Segregation in Evolving Social Networks. *Proceedings of the National Academy of Sciences*, 108(21): 8605-10.

Jugert, P., Leszczensky, L. and S. Pink. (2018). The Effects of Ethnic Minority Adolescents' Ethnic Self-Identification on Friendship Selection. *Journal of Research on Adolescence.* 28(2). 379-95.

Kalter, F., Heath, A.F., Hewstone, M. Jonsson, J.J., Kalmijn, M. Kogan. I., and F. van Tubergen. (2016) Children of Immigrants Longitudinal Survey in Four European Countries (CILS4EU) – Full version. Data file for on-site use. GESIS Data Archive, Cologne, ZA5353 Data file Version 1.2.0, doi:10.4232/cils4eu.5353.1.2.0.

Frank Kalter, Jan O. Jonsson, Frank van Tubergen, and Anthony Heath (eds.). 2018. *Growing Up in Diverse Societies: The Integration of the Children of Immigrants*

in England, Germany, the Netherlands, and Sweden. Oxford: Oxford University Press.

Koskinen, J. and T.A.B. Snijders (2019). Random-Effect Multilevel Stochastic Actor-Oriented Models (*in preparation*).

Kreager, D. A. (2008). Guarded Borders: Adolescent Interracial Romance and Peer Trouble at School. *Social Forces.* 87(2). 887-910.

Koskinen, J.H. And T.A.B. Snijders. (2007): Bayesian inference for dynamic social network data. *Journal of Statistical Planning and Inference* 137 (12), 3930–3938.

Leszczensky, L. (2018). Young Immigrants' Host Country Identification and their Friendships with Natives: Does Relative Group Size Matter?. *Social Science Research.* 70. 163-75.

Leszczensky, L., and S. Pink. (2015). Ethnic Segregation of Friendship Networks in School: Testing a Rational-Choice Argument of Differences in Ethnic Homophily between Classroom- and Grade-Level Networks. *Social Networks.* 42. 18-26.

Lusher, D., Koskinen, J. and Robins, G. (2010) Exponential Random Graph Models for Social Networks: Theories, Methods and Applications. Cambridge: Cambridge University Press.

Macy MW, Willer R. (2002). From factors to actors: Computational sociology and agent-based modeling. *Annual Review of Sociology.* 28.143-66.

McPherson, M., Smith-Lovin, L., and M.J. Cook. (2001). Birds of a feather: Homophily in social networks. *Annual review of sociology.* 415-44.

Moody, J. (2001). Race, School Integration, and Friendship Segregation in America. *American Journal of Sociology.* 107(3). 679–716.

Munniksma, A., Flache, A., Verkuyten, M., and R. Veenstra. (2012). Parental Acceptance of Children's Intimate Ethnic Out-Group Relations: The Role of Culture, Status, and Family Reputation. International Journal of Intercultural Relations. 36(4). 575-85.

Munniksma, A., Stark, T.H., Verkuyten, M., Flache, A., and R. Veenstra (2013) Extended intergroup friendships within social settings: The moderating role of initial out-group attitudes. *Group Processes & Intergroup Relations.* 16(6). 752–70.

Munniksma, A., Verkuyten, M., Flache, A., Stark, T.H., and R. Veenstra. (2015) Friendships and out-group attitudes among ethnic minority youth: The mediating role of ethnic and host society identification. *International Journal of Intercultural Relations.* 44. 88–99.

Pinter, B. and A.G. Greenwald, (2011). A comparison of minimal group induction procedures. *Group Processes & Intergroup Relations.* 14 (1). 81–98.

Prell, C., and Y.-J. Lo. (2016). Network formation and knowledge gains. *The Journal of*

Mathematical Sociology. 40.

Ripley, R.M., Snijders, T.A.B., Boda, Z., Vörös, A., and P. Preciado (2018). Manual for SIENA version 4.0. Oxford: University of Oxford, Department of Statistics; Nuffield College.

Rusbult, C. E. and Buunk, B. P. (1993). Commitment processes in close relationships: An interdependence analysis. *Journal of Social and Personal Relationships.* 10(2). 175-204.

Scheepers, D., Spears, R., Doosje, B., and A.S.R Manstead. (2006). The social functions of in-group bias: Creating, confirming, or changing social reality. *European Review of Social Psychology* 17 (1), 359–96.

Schelling, Thomas C. 1978. Micromotives and macrobehavior. WW Norton & Company.

Smith, S., Maas, I. and F. van Tubergen. (2014). Ethnic ingroup friendships in schools: Testing the by-product hypothesis in England, Germany, the Netherlands and Sweden, *Social Networks.* 39. 33-45.

Smith, S., Van Tubergen F., Maas, I. and D.A. McFarland (2016). Ethnic composition and friendship segregation: Differential effects for adolescent natives and immigrants. *American Journal of Sociology.* 121 (4). 1223-72.

Snijders, T. A. B. (2001). The statistical evaluation of social network dynamics. Sociological methodology. 31(1). 361-95.

Snijders, T. A. B., Van de Bunt, G. G., and Steglich, C. E. G. (2010). Introduction to stochastic actor-based models for network dynamics. Social networks, 32(1):44-60.

Snijders, T. A. B. and C. Steglich. (2015). Representing Micro–Macro Linkages by Actor-based Dynamic Network Models. *Sociological Methods & Research* (44). 222–71.

Stadtfeld, C., Snijders, T.A.B., Steglich, C., and M.A.J. van Duijn. (2018). Statistical Power in Longitudinal Network Studies. *Sociological Methods & Research.* forthcoming.

Stark, T. H. (2015). Understanding the Selection Bias - Social Network Processes and the Effect of Prejudice on the Avoidance of Outgroup Friends. *Social Psychology Quarterly.* 78 (2). 127-50.

Stark, T. H., Flache, A., and Veenstra, R. (2013). Generalization of Positive and Negative Attitudes Toward Individuals to Outgroup Attitudes. *Personality and Social Psychology Bulletin*, 39(5). 608-22.

Steglich, C., Snijders, T. A., and Pearson, M. (2010a). Dynamic networks and behavior: Separating selection from influence. *Sociological Methodology.* 40(1). 329-93.

Steglich, C., Snijders, T. A. B., and Pearson, M. (2010b). 8. Dynamic Networks and Behavior: Separating Selection from Inuence. Sociological Methodology,

40(1):329{393.

Tajfel, H. (1974). Social identity and intergroup behaviour. *Social Science Information/sur les sciences sociales.* 13(2). 65-93.

Tajfel, H. and Turner, J. C. (1979). An integrative theory of intergroup conflict. *The social psychology of intergroup relations.* 33:47.

Turner, R. N., Hewstone, M., and Voci, A. (2007). Reducing explicit and implicit outgroup prejudice via direct and extended contact: The mediating role of self-disclosure and intergroup anxiety. Journal of personality and social psychology. 93(3). 369.

Uehara, E. (1990): Dual Exchange Theory, Social Networks, and Informal Social Support. *American Journal of Sociology.* 96 (3). 521–557.

Verbrugge, L.M. (1977). The Structure of Adult Friendship Choices. *Social Forces* 56 (2), 576–97.

Vermeij, L., van Duijn, M.A.J., Baerveldt, C, (2009). Ethnic segregation in context: Social discrimination among native Dutch pupils and their ethnic minority classmates. *Social Networks* 31, 230–39.

2 CHOICE AND STRUCTURAL HOMOPHILY FOR NATIVE AND NON-NATIVE PUPILS IN CLASSROOM FRIENDSHIP NETWORK

Co-Author:

Dr. Enis Bicer, Akkon-Hochschule für Humanwissenschaften, Berlin

Email: enis.bicer@akkon-hochschule.de

2.1 Introduction

Social relations are important and central aspects of human life. Friendships can satisfy human needs like the desire for social acceptance and belonging (e.g. Baumeister and Leary, 1995; Denissen et al., 2008), provide emotional or financial support as well as advice about various topics (e.g. see Furman and Buhrmester, 1992; Hartup and Stevens, 1997; Helsen et al., 2000; Uehara, 1990). Moreover, friend groups constitute an important reference for one's own identity (e.g. Meeus and Dekovic, 1995; Weeks and Pasupathi, 2010).

When looking at the question of who bonds with whom, most social relations reflect the saying "birds of a feather flock together" (see McPherson et al., 2001). This describes the tendency of individuals to form relationships based on their similarity to others. Consequently, various social relationships (for example peer groups) are overly similar with regard to specific attributes, like race/ethnicity, gender, socioeconomic status, opinions and beliefs, behaviour, and so on. From the perspective of social networks this leads to the common finding of segregated clusters. Friendship networks within schools,

for example, show a clear tendency for gender and race/ethnic segregation (e.g. Lubbers et al., 2011; Moody, 2001).

While this is an established finding in social networks, the question about the particular mechanisms (and their interplay) causing this pattern is still not a widely and extensively explored issue (Lewis, 2015: 6). While Moody (2009) describes *node* and *edge-based* mechanisms as theoretical roots for segregation in networks, until now only a few empirical studies have looked in depth at causes of segregation by acknowledging both mechanisms at once (see Goodreau et al., 2009; Grund and Densley, 2015; Kossinets and Watts, 2009; Wimmer and Lewis, 2010). Although these studies indicate an interplay between preference and structural based processes of segregation, they either do not specifically account for important factors like race and ethnicity or only apply cross-sectional data making causal inference about the evolution of segregation in networks rather difficult.

This study is motivated to do an analysis of processes explaining the dynamic evolution of networks regarding ethnic segregation. Specifically, we want to examine the role of structural network embeddedness, reflected in transitivity, for the segregation of friendship networks between native and non-native German pupils. Our analysis extends the investigation of general transitivity by differentiating transitive closure to occur within or between both groups. This allows to get a better understanding about the role of homophily on a dyadic and triadic level.

The dataset we use provides directed friendship networks in German school classes, longitudinally collected in three waves. We categorize participating pupils as native or non-native Germans and aim to explore the evolution of their within- and between-group friendship ties. By referring to ideas based on social psychology and intergroup research, theoretical expectations are formed about dyadic and triadic homophily

processes for both groups. The subsequent analysis starts by promoting descriptive statistics about the networks we look at. We then apply Bayesian hierarchical stochastic actor oriented models (HSAOM) (Koskinen and Snijders, 2018) to investigate group differences between natives and non-natives.

This study contributes by investigating causes of segregation by applying a longitudinal perspective. Second, by differentiating for node-specific attributional composition in triplets, and simultaneously controlling for dyadic homophily, this study investigates how both mechanisms potentially contribute to segregated friendship networks.

2.2 Theoretical background

2.2.1 Node-based mechanisms of segregation

Building up a personal relationship can be regarded as a (more or less) deliberate action of every ego in a network choosing an 'appropriate' network alter to bond with. Therefore, it is logical to locate sources of tie creation on an *actor*- or *node*-level (see Moody, 2009: 449). Actors often show a preferential tendency to create new relations based on node-attribute similarity, which is called *choice homophily* (Lazarsfeld and Merton, 1954; McPherson and Smith-Lovin, 1987), and has been extensively examined and well-documented in a wide range of studies (e.g. Baerveldt et al., 2004; Hallinan and Williams, 1989; McPherson et al, 2001; Smith et al., 2014; Verbrugge, 1977; Vermeij et al., 2009). This *assortative* process of tie evolution (Rivera at al., 2010: 94) is a *node-level* mechanism which can contribute to segregation in networks. Regarding ethnicity, individuals' preferences for same ethnic friends contribute to a higher chance for same ethnic dyads to be directly connected over time, thus segregating networks by ethnic categories.[6]

6 It can be argued that homophily is multidimensional (e.g. similarity with the combination of two dimension; gender and ethnicity) and it has been found that being similar on a second dimensions has not as strong an additional effect as being similar in one dimension compared to not being similar with regard to both (Block and Grund, 2014). We argue that for the purpose of our work it is sufficient to

Friendships form in a process of repeated interactions between individuals. With each of these consecutive interactions that is perceived as positive by an individual, their opinion of or sympathy for the other will improve. In other words, individuals will expect their investment of time that is spend with the other to be worthwhile, such that eventually one actively seeks these interactions. After each consecutive interaction, individuals update their expectations about the outcomes of spending time with another. If the perception of a relationship to an alter reaches an individual's subjective threshold, they will start to label it as a friendship.

On the other hand, due to its informal and voluntary characteristics, friendship relations are also prone to interpersonal exploitation. Building up a friendship with someone new implies interpersonal sympathy or attraction (Fehr, 2008; Sprecher and Felmlee, 2008), as well as a minimum of trust in the other person´s sincerity and reliability. When sharing private experiences, thoughts and emotions with a close friend or asking for advice in private life matters, we do not necessarily want those things to become public. Therefore, it is important, that the friend one confides in is perceived to be a trustworthy and discreet person; someone unlikely to misuse entrusted shared private information. Considering this, inceptive self-disclosure in the process of emerging relations (e.g. Derlega et al., 2008; Duck and Miell, 1986; Levesque et al., 2002; Miell and Duck, 1986) often goes along with some personal risks of disappointment, emotional hurt, or feelings of having been deceived when the assumed (new) friend turns out not to be trustworthy. Consequently, the process of building up a friendship with an unknown person is accompanied by several uncertainties (see e.g. Ayres, 1979; Berger and Calabrese, 1975). It requires the will to make some investment into a relationship with a stranger, while the success of this advanced input is not guaranteed. In the outlined

focus on one of the two dimensions while controlling for the other and investigate homophily of ethnic groups.

approach about individuals becoming friends with another we do not assume that they are aware about all expected costs and make a rational decision about who they are friends with. Rather, it is an analytical approach that captures some of the factors influencing individuals' everyday choices about whom to spend more time interacting with.

We argue that one reason to prefer same ethnic friends comes from higher uncertainties about the positive outcome of (cognitive and emotional) investment into a respective outgroup alter. There is almost consistent evidence that attitudes and behaviours towards people differ considerably when they are from the same or a different group. This includes favoring in-group members by beneficial resource allocations, and derogating members of the respective out-group (e.g. Bettencourt et al., 2001; Pinter and Greenwald, 2011; Scheepers et al., 2006; Tajfel, 1970). Moreover, stereotype beliefs, prejudices and preferences for social distance and avoidance work as strategies contributing to intergroup boundary solidification and self-concept preservation (see e.g. Fein and Spencer, 1997). Regarding intergroup relations in ethnically diverse societies, research on the rejection of people with immigrant background or forms of hostility towards foreigners, are prominent examples for burdened interethnic group relations (e.g. Semyonov et al., 2006; Schneider, 2008; Wagner et al., 2008). Leaning on the prominent theory of social identity (Tajfel and Turner, 1986), categorizing the social world into “we” and “them” plays an important role, defining the situation in a competitive or overly confrontational way. This often is accompanied by various types of negative feelings and behaviour towards members of an out-group, e.g. people with immigrant background.

Consequently, we conclude that the specific conditions for realizing between-group friendships must overcome two obstacles. First, the aforementioned uncertainty

inherently accompanying the process of friendship initiation in general. Here the basis of uncertainty about positive returns from interactions and the potential friend´s trustworthiness is due to unknown features of his or her individual personality. This is true for almost every interactive situation where two strangers become acquainted with each other and is independent of the group affiliations of the involved actors. We term this a situation of *interindividual uncertainty.* The second obstacle becomes specifically relevant to intergroup relations and refers to the burden of stereotypical beliefs about a person's characteristics due to his/her perceived affiliation to an out-group. Research indeed shows that others are frequently seen and evaluated based on group-related expectations and less often on their actual individual personal features (e.g. Darley and Gross, 1983). These beliefs are often negatively connoted, for example reflecting doubts about the trustworthiness of the other. Various experimental and other empirical studies demonstrate lower levels of trust to out-group members and adverse consequences on cross-group cooperation (e.g. Buchan et al., 2002; Çelebi et al., 2014; Insko and Schopler, 1998; Macy and Skvoretz, 1998; Yamagishi and Kiyonari, 2000; Yuki et al., 2005).[7]

Thus, we conclude that building up a between-group relationship between natives and non-natives on the one hand will face general interindividual difficulties regarding the time one enjoys with a potential friend as well as their reliability and trustworthiness. On the other hand, it will also be burdened with specific negatively biased expectations about the others character traits and future behavior based on perceived out-group affiliations. We term the latter as *intergroup uncertainty* which is a supplementary adverse condition specifically relevant for potential between-group friendships. Therefore, we expect in-group friendship initiation being perceived as more favorable

7 Beside the specific intergroup case the study of Prisbell and Andersen (1980) also demonstrates an association between perceived homophiles in attitudes and values with reduced uncertainties in interaction.

and less 'risky' in comparison to cross-group relations and consequently will be preferred by individuals. Furthermore, notable studies are supporting the supposed associations between peer group contact and friendship networks with out-group attitudes (e.g. Munniksma et al., 2013; Munniksma et al., 2015; Stark et al., 2013; Stark, 2015). Regarding our theoretical thoughts from above, we thus are following that choice homophily as a node-based mechanism makes native and non-native within-group friendships more likely than between those groups (**H1a**).

Regarding the composition of the non-native population in Germany one has to consider, that it is a diverse category. In the 60s the so called “guest workers” have been the first significant non-German group to immigrate to the country, coming mainly from Turkey, Italy, Greece, Spain and the former Yugoslavian Republic. Since then, and especially the end of the Cold War era, immigration has become more and more diversified, for example by receiving asylum seekers from within and outside of Europe (see Bade and Oltmer, 2010). Consequently, the immigrant or non-native population living in Germany today consists of persons with different migration and ethnic backgrounds. While the group of natives can identify and differentiate itself from non-natives (irrespective or their ethnic origins), it stands to questions if non-natives identify as such in difference to the native population (namely as non-natives) or rather identify with the group of their respective origin, e.g. the Turkish or Russian immigrant group. We have to acknowledge for this and therefore it is worth to note, that natives' between-group ties would be stronger subject to group identification in line with this differentiation than is the case for non-natives. For this reason we see the necessity of refining the above hypothesis by taking into account possible group differences between natives and non-natives.[8] Thus, pupils from the group of natives more strongly identify

8 Wimmer and Lewis (2010) are emphasizing the potential role of group-specific differences in network endogenous mechanisms like reciprocation and balancing. While they cannot find higher probabilities for intraracial reciprocity, they have not been able to test for group-specific triadic closure effects, due

as such than immigrants do as non-natives. Therefore, we expect natives to have in average a stronger dyadic tendency (choice homophily) to create and maintain in-group friendship ties compared to natives (**H1b**). Of course, this does not account for possible differences between non-natives of different origin groups and their respective group identifications. Testing the here made assumption that natives more strongly identify as such while non-natives rather hold ethnically diverse group identifications could be attempted in a first step by linking non-native homophily to the diversity of ethnic origin groups in the composition of classrooms. However, this differentiation would go beyond the aim of this work to firstly show differences between the group of natives and non-natives and secondly investigate the role of network structures (in this case transitivity) and homophily.

2.2.2 Edge-based mechanisms of ethnic segregation

Besides node-based mechanisms of network evolution, the social environment or context an actor is embedded in also plays an important role. This refers to structural mechanisms, explaining the creation of ties between two actors *i* and *j* due to relations with or between others in the network surrounding the respective dyad *ij*. These *relational mechanisms* (Rivera et al., 2010: 97) of network evolution can be subsumed as *edge-based* models, because they are specifically interested in "how edges emerge or dissolve as a function of surrounding edges" (Moody, 2009: 450). One of the most established *local forces* (Agneessens and Skvoretz, 2012: 1523) for tie creation is the tendency for triads to show closure over time.[9] Based on the classical theory of balance (Cartwright and Harary, 1956; Heider, 1946) and research on various structural

to the unavailability of race-specific transitivity effects in ERGM´s at the time the study have been conducted.

9 Beside triadic closure, the most prominent relational mechanism is the tendency for ties to be reciprocated. As our explicit interest of this research is not on reciprocity we will not discuss reciprocity here in further detail. Nevertheless it`s undoubtedly important and in the forthcoming empirical analysis controlled for.

tendencies of networks (e.g. Davis, 1967; Davis and Leinhardt, 1972) the concept of *transitivity* is considered as the most generalized mechanism of triadic closure and network structure (Holland and Leinhardt, 1970; 1971). Transitivity occurs in situations where two actors *i* and *j* are indirectly connected via an intermediate actor *k* and subsequently create an edge (*i* to *j)*. This reflects the concept of "friends of friends becoming friends". Transitive closure contributes to the formation of local clusters in networks and found empirical confirmation in a variety of studies, thus is considered a regularity of network formation (e.g. Baerveldt et al., 2008; Dijkstra et al., 2011; Friemel, 2012; Hammer, 1979; Kossinets and Watts, 2009; Lubbers et al., 2010; Wimmer and Lewis, 2010).

Regarding mentioned uncertainties in friendship making, being indirectly bonded to an alter via a 2-path enhances the chances for direct ties in future for several reasons. First, by posing an opportunity structure to get personally acquainted with others. Second, by signaling that an already related alter took the risk and succeeded. Furthermore, a shared third party importantly contributes to the emergence and amplification of trust between not directly connected persons (see e.g. Burt and Knez, 1995; Coleman, 1990: 180ff.). In line with this, uncertainty reduction theory (Berger and Calabrese, 1975; Berger and Bradac, 1982) states that people seek to reduce uncertainties by asking a third party for information about a yet unknown target person (see also Berger and Kellermann, 1994). Thus the transmission of relevant information[10] about an alter´s trustworthiness contributes to the emergence of interpersonal trust between an ego and a respective alter (e.g. Barrera, 2005; Barrera and Buskens, 2009; Buskens, 2002). Concluding from this, we expect intermediate actors to enhance the chances for ties to be embedded in triads (**H2**).

10 The idea of information flowing through indirect ties between not directly connected actors is related to the conceptualization of brokerage in exchange networks by Marsden (1982).

One major contribution of this study is to show that node- and edge-based mechanisms have an independent as well an interrelated contribution to the formation of homogenous relational patterns in networks and thus leading to segregation. As has been pointed out by Grund (2014), networks with a basic tendency for homophilous tie creation have a higher chance to indirectly connect more actors of an in-group than an out-group. Therefore, part of the homophily process leading to segregation occurs as a structural effect of transitive closure rather than preference. Figure 2.1. illustrates transitivity for different compositions of transitive triplets. The upper part a) shows compositions where ties exist between individuals *i* and *h* and *h* and *j* (indicated by solid lines) and a tie that can be realised between individuals *i* and *j* of the same background (dashed lined). The right hand side shows the case a1) where the intermediate actor *h* has the same background as actors *i* and *j* and a2) where the intermediate actor *h* has a different background than actors *i* and *j.* The lower part b) reflects the same logic but for actors *i* and *j* having a different background. If a network starts to form and individuals have a slight preference to create in-group ties (dyadic homophily), this will lead to more two-paths between individuals of the same background. With a general tendency of individuals to close these two-paths transitively at subsequent time points, this results in more in-group transitivity such as occurring in all-in-group triads as depicted in a1).

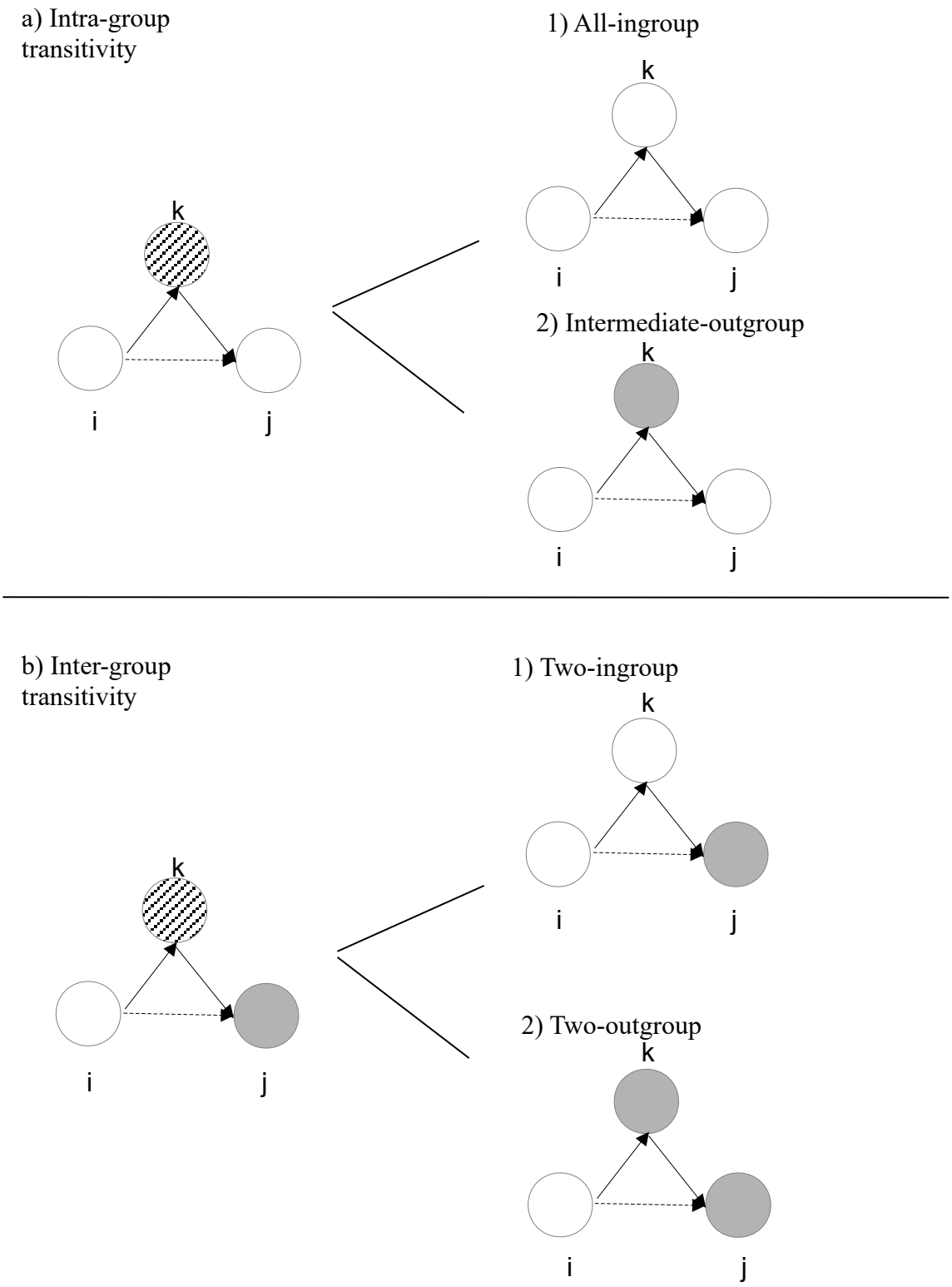


Figure 2.1 : Composition of transitive triplets

This can be traced back to *node-* as well as to *edge-based* mechanisms for the within-category tie (*ij*) to emerge. On the one hand, actor *i* has an inclination for in-group friends and consequently chooses in-group person *j* as a friend (choice homophily - *assortative mechanism*). On the other hand, the tendency to befriend a friend´s friend may also be responsible for *i* befriending *j,* independently of particular group associations (transitivity - *relational mechanism*). Moreover, it is conceivable that both mechanisms jointly play their part in contributing to the within-group tie (*ij*) to develop or even mutually influence each other, for example by an interactive amplification process. This demonstrates the difficulty in distinguishing both mechanisms theoretically and underlines the necessity of simultaneously accounting for both in an empirical analysis of evolving networks. This has been underlined by previous empirical studies focusing on the interplay of *edge-* and *node-based* mechanisms. Kossinets and Watts (2009) show that the effect of pairwise similarity on the probabilities for tie formation is partly mediated by structural proximities, like shared acquaintances. The study of Goodreau et al. (2009) concludes that both mechanisms are responsible for mutual friendships by grade, sex and (most) race groups. In the study of Grund and Densley (2015) closure mechanism for same-ethnic triads is specified. From their corresponding ERGM analyses, they conclude that the effect of ethnic homogeneity in dyads loses much of its relevance when controlling for higher-level mechanisms, namely ethnic homogenous triadic closure. Finally, in their cross-sectional study Wimmer and Lewis (2010) demonstrate that racial homogeneity of college students' Facebook networks is not only based on mechanisms of racial homophily, but also (and even to a higher extent) stems from edge-based mechanisms like reciprocity and transitivity. While transitivity can amplify the contribution of dyadic homophily for

the segregation of a network, this does not imply a tendency for actors to rather have embedded ties to the in-group than the out-group (*structural homophily*). Thus if *structural homophily* (or it's counterpart *heterophily)* occurs not only would it further amplify tendencies for segregation, not accounting for this can also lead to a biased estimation of dyadic homophily. While homophily on the dyadic level can be understood straight forward as the tendency of individuals to rather have in-group than out-group relations, this is more complicated for the triadic case. To systematically separate the mechanisms of structural embeddedness and homophily for both groups, we first look at the role of transitivity for intra- and inter-group ties compared to not-embedded dyads separately for the group of natives and non-natives separately. Following, we formulate our expectations about group differences in structural homophily for both groups.

For embedded within-group friendships of natives, we assume potential risks and uncertainties to be solely located on an interindividual level (*interindividual uncertainty*), meaning that an intermediate person transmits relevant interindividual information about each other through the 2-path connection between ego and alter. This makes natives' in-group ties more likely when embedded in a transitive triplet. Because in-group ties for the group of non-natives can also connect individuals of different ethnic origin (as for example pupils with Russian and Turkish background), intermediate actors can also contribute to reduce *intergroup uncertainties*. In terms of rational choice arguments such as made by Lescczensky and Pink (2015) or Block (2015) relationships between individuals of the same group comes with lower costs than relationships between individuals of different groups. On the one hand, transitive triplets can be considered as a group, providing lower costs to individuals than not-embedded dyads. Secondly, pupils having native or non-native backgrounds can

constitute an additional grouping differentiation. Following the arguments made for group differences of dyadic homophily, natives might hold a stronger identification with their group of being natives, than non-natives do. Along these lines of reasoning, intra-group transitivity for natives is more likely to provide a double group benefit of transitive embeddedness and identification with the group of natives. Therefore, it can be expected that *intragroup transitivity* is stronger role for natives than for non-natives (**H3a**).

Regarding between-group ties, the intermediate actor reduces *interindividual* as well as *intergroup uncertainties*, irrespective of the group an actor belongs to. Inter-group transitivity provides a lower cost situation than inter-group dyads do, due to the embeddedness through the intermediate actor. For this case, natives would not benefit from additionally group identification. Thus, while transitivity makes embedded between-group relations more attractive (or less costly) than between-group dyads, there is no reason to expect this to differ between groups.

From this we can derive our expectations about structural homophily and how it differs between groups. Similar to dyadic homophily, structural homophily (heterophily) would be given if a group has a larger (smaller) tendency for intragroup transitivity then for intergroup transitivity. Since we argue that intragroup transitivity is stronger for natives than for non-natives, while inter-group transitivity does not differ, it follows that natives should show stronger differences between intra- and intergroup transitivity compared to non-natives. Thus, structural homophily is expected to be stronger for natives than for non-natives (**H3b**).[11]

11 By similarly accounting for group affiliations of actors in triads, Gould and Fernandez (1989) specify five distinct brokerage types in transactional networks. Nevertheless in their work they do not formulate explicit theoretical expectations about triadic closure mechanisms, which in contrast constitute the central aspect of our study. Therefore we do not apply the brokerage approach of Gould and Fernandez (1989) in our theoretical frame, but instead rely on balance theory and transitive closure in combination with a micro level explanation on the basis of intergroup relations theory.

2.3 Data and method

The empirical data for our analyses is from the longitudinal study "Social Networks and Social Integration of Schoolchildren" funded by the German Research Association (DFG).[12] It was conducted in two German cities, namely Bremen and Bremerhaven, located in the north of the country with both having a considerable proportion of inhabitants of non-German family origin. The used dataset provides information about full social networks of 22 classrooms which participated in three consecutive years (2010-2012). The first wave was collected about half a year after the classrooms were set together in the 5th grade, which is the first year of secondary school in Germany. The participating pupils' age is between 10 and 14 years, and the classrooms consist of 20 to 35 pupils. In each wave, participating pupils filled in a battery of sociometric measures. Next to other network dimensions, the participants were asked to name other pupils within the same classroom whom they consider to be their friends, who they invite to their birthday and who they spend time with on the school-yard. A friendship tie was considered to be present when a pupil nominated another on all three network dimensions. Following this, we used these classroom-based directed friendship networks as the dependent variable in our analysis. Central to this analysis is the construction of the variable indicating whether pupils are of German descent ("native") or not ("non-native"). We coded pupils as native when both parents are of German origin, and pupils of whom at least one parent is not of German origin have been coded as non-native. Additionally, we consider the sex of the participants ("female") to account for the well-known tendency of young adolescents to befriend pupils of their own sex.

First, we will investigate the basic descriptives of our networks. This includes the distributions of size of female and non-native students in the classrooms. We will also

12 DFG grant no. WI 3423/1-1, AOBJ 574807

look at the distributional measures of degrees, missings and densities for our friendship networks throughout all three waves. Furthermore we will present overall segregation levels of the classroom-networks by applying two different suitable indices. In addition, we will present descriptive statistics for tie creation in dyadic and triadic network structures while taking their specific group composition into account. For the next step we will apply (multilevel) stochastic actor oriented models (SAOM) (Snijders et al., 2010b) as an appropriate way to analyse changes of our longitudinal networks. The SAOM models the network evolution in continuous-time, assuming that actors make changes to their ties at different points in time. The main target of inference is to model these changes, and the basic idea is to model the choices of ego *i* in so-called *mini-steps*. For each of these mini-steps an individual chooses whether a new tie will be created or an existing tie will be dissolved.

The mini-steps are rarely directly measured in existing datasets, meaning that the exact sequence of decisions in between waves is unobserved. The estimation of model parameters can still be conducted by applying any of three estimation procedures that account for this missing information: method of moments (Snijders, 2001); Bayesian (Koskinen and Snijders, 2007), and maximum likelihood (Snijders et al., 2010a). In order to distinguish two structural equivalent types of transitivity, we need to estimate the model using one of the two likelihood-based frameworks. Furthermore, as we have repeated observations on multiple independent networks (the classrooms) we need to employ a hierarchical model that can handle heterogeneity between classrooms. With this aim, we apply the recently developed Bayesian Hierarchical SAOM (HSAOM) (Koskinen and Snijders 2018). The HSAOM has the benefit of allowing us to fit the same model specifications to all classrooms, something that previous analyses showed was not possible due to, among other things, the presence of some small classrooms. It

is conceptually similar to classical regression-type multilevel models with regard to differentiating variation on the individual and on a group level. In the HSAOM approach, each classroom g=1,...,G is assumed to follow the same SAOM but with a classroom-specific set of parameters $\theta^{[g]}$. These parameters in turn are assumed to be drawn from a population-level model which is a multivariate normal distribution with a mean μ and covariance σ. The target of inference is primarily to estimate the population mean μ rather than the classroom-level parameters $\theta^{[g]}$. The other parameters are assumed to be constant across classrooms.

The estimation procedure involves drawing samples of three key posterior variates: the population parameters, the classroom-level parameters, and the sequences of unobserved changes in-between waves. These variates are generated through an iterative Markov-chain Monte Carlo (MCMC) sampling scheme. Inference is based on these draws from the posterior values given the observed data. Because inference is based on samples from the posterior distributions, one needs to be careful to ascertain that the MCMC samples efficiently; first models have been run with 300 main phase iterations (default options), followed by a rerun of 1000 iterations during the main phase, and final estimations are run with another 5000 iterations (Ripley et al., 2018) using flat priors. From a non-Bayesian perspective, the posterior mean and standard deviation of parameters can be interpreted as approximations to the MLE and standard error, respectively. When calculating linear combinations or differences in effects, these calculations are done for each iteration separately, such that the mean and standard deviation of the respective calculated values again MLE approximations and standard errors. If one wishes to calculate significance, this can simply be done by calculating the proportion of draws which are above or below zero.

Since the main phase estimations are likelihood based, we are able to differentiate

transitivity of tie creation between (1) actor i and j and (2) actor i and k. Because our theoretical outline aims at the first situation, only this will be further specified for different group compositions. The second aspect will only be included in the model as a general control effect, allowing us to get clearer estimates of the network situation we are interested in. Also, as our research interest is not focused on between-classroom variation of effects (i.e. homophily), nor on explanations of that variance, we force the classroom-level parameters[13] to be the same $\theta^{[g]}=\theta^{[h]}$ for all g and h[14] and do not include classroom characteristics (i.e. second level variables) as predictors. Though this would provide further interesting insights, it is beyond the scope of this paper. All SAOM applications are carried out using the RSiena Software (Ripley et al., 2018).

2.4 Results

2.4.1 Descriptives

Table 2.1 shows that over all waves the analyzed classes consist of an average of 27 pupils and of 53% females. The proportions of non-natives in the classrooms range between 7% and 54% with an average of 43%. This variation in the composition of the classrooms challenges the model assumption that homophily mechanisms are the same over all classrooms. Especially in classrooms with a very low proportion of non-native pupils this results in only one or two non-natives being present. Of course, this restricts the opportunity structure for in- or out-group friendships considerably, even though these classrooms only constitute a small proportion of our sample. The implications of this for our findings are further discussed in the conclusions section.

With regard to the observed friendship network descriptives, it can be seen that the average degrees increase throughout all three waves and the networks become denser,

13 Except for outdegree
14 While g represents the overall parameters, h represents the individual classroom parameters.

meaning that more ties are created than dissolved. This difference is especially prominent between the first and the second wave where the average outdegree increases by more than 1.5, but only by 0.3 between the second and third wave. Thus, in the first period on average each pupil nominates more than one additional friend, while in the second period only each third pupils nominates an additional friend. In line with the fact that our data was collected only half a year after pupils have been assembled into classrooms this opens questions about time heterogeneity in the mechanisms driving the network evolution. During the first half year of the constituted classrooms, a network first has to emerge, i.e. friendship ties have to be established such that mechanisms of friendship creation will be more prominent that those of friendship maintenance. The stronger increase of the outdegree between the first two observations can suggest that this might still be the case to some extent. This is in contrast to the time period between the second and third observation where tie maintenance might be more relevant. Accounting for such time heterogeneity of network mechanisms is especially important when investigating mechanisms of network creation and maintenance separately. While this would go beyond the scope of this work, it still can have implications for the findings which will be discussed in the conclusion section.

It can be also seen that, for the second time point, the range and standard deviation of the degrees between classrooms becomes larger. This might indicate that the chosen HSAOM method is appropriate, since it does not assume the respective effects that model overall changes of the networks (i.e. rate parameters and outdegree) to be constant between classrooms.

General	min	mean	max	sd
N pupils	20	27	35	4.12
% female	.30	.53	.68	.10
% non-native	.07	.43	.54	.15

Friendship Network	Wave 1				Wave 2				Wave 3			
	min	mean	max	sd	min	mean	max	sd	min	mean	max	sd
Degree	2.11	5.17	8.66	1.62	3.56	6.83	13.80	2.32	5.06	6.92	11.79	1.58
Missings	.00	.14	.44	.11	.00	.15	.33	.09	.00	.13	.48	.12
Density	.11	.20	.31	.05	.16	.27	.53	.09	.15	.27	.49	.07

N classes = 33

Table 2.1: Basic descriptives

Assessing the overall levels of ethnic segregation in the classrooms, for each time point of measurement we look at the overall distribution of the (reversed) E-I index (Krackhardt and Stern, 1988) and Yule's Q (Yule, 1912) (see figure 1.1). Higher values in the inverse E-I index mean that more in-group than between-group ties exist compared to all existing ties. This implies that it is sensitive to sizes of the groups. On the other hand, Yule's Q is not sensitive to group sizes (Borgatti et al., 2013: 274), since it compares existing in-group ties by not existing between-group ties minus existing between-group ties by not existing in-group ties. As figure 2.2 shows, both indexes show moderate levels of segregation. Throughout all three measurement points the magnitude of the E-I index stays rather constant. With regard to Yule's Q, the overall segregation levels are generally somewhat lower, though still apparent for the majority of classrooms.

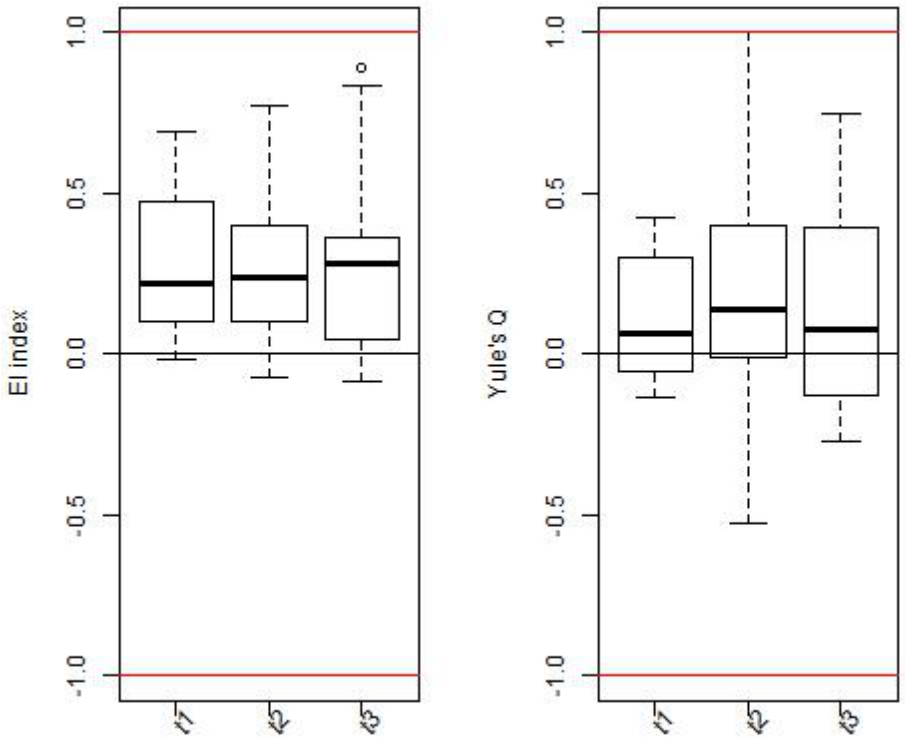


Figure 2.2: Measures of segregation at each wave

To gain first insights about group difference in segregation, figure 2.3 illustrates the proportion of realised ties within and between groups, separately for natives and non-natives at the three waves. For each wave (t1, t2 and t3) a set of box-plots represent the distribution of proportions of existing ties from possibly existing ties that were calculated within and between each group separately. The first graph for t1 shows on the left side the proportion of in-group ties realised by non-natives in grey, and of between-group ties in white. Next to it are the findings for the native group. Inspecting these results for the first wave shows that non-natives seem to have an even higher proportion of realised ties to natives than to other non-natives, while natives show a higher proportion of within-group ties. In other words, at the first wave, natives made more

friends with each other, whereas non-natives preferred natives. However, from the second wave onwards this changes such that non-natives become more inclined to have more within-group than between-group ties, too. Firstly, this underlines that mechanisms of homophily differ between both groups. Secondly, it suggests that, in order to understand the underlying phenomenon of these findings, it might be important to further investigate differences of tie creation and maintenance. This is because the observed change in the proportion of realised in- and out-group friendships of non-natives could be driven by non-natives to create many more in-group friendships, dissolve several of their friendship nominations to natives or a combination of both. Thirdly, the differences of realised between-group friendship between natives and non-natives might be related to reciprocity. At the first wave, non-natives might nominate more natives than other non-natives, but aren't nominated back (i.e. these nominations are not reciprocated). In relation to the previous point, this might result in a dissolution of such one-sided friendships from non-natives. For understanding the role of transitivity for mechanisms of homophily, one can followe the argumentation of Block (2015) that dyads which are embedded in triplets would not necessarily need to be reciprocated to be maintained. Thus, the findings presented in figure 2.3 can suggest that one-sided ties from non-natives to natives are not be embedded in triads, or if they are dissolve more often than those embedded within groups. Thus, in the next step we investigate the census of transitive triplets, two-paths and transitive closure for intra- and inter-group ties to gain first insights about trends for structural homophily.

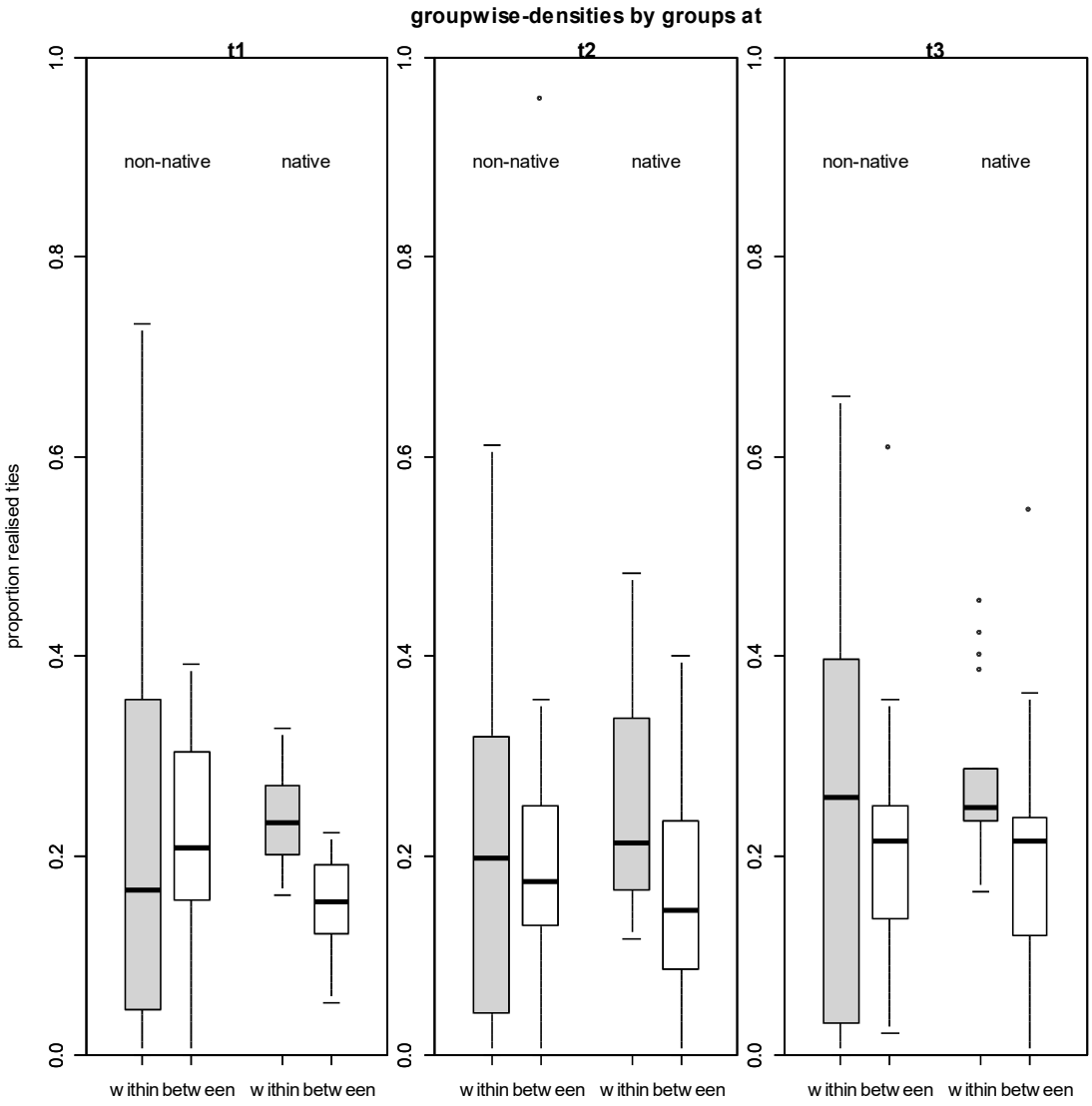


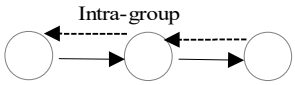
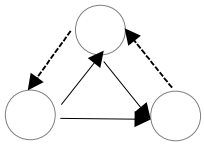
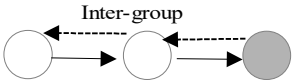
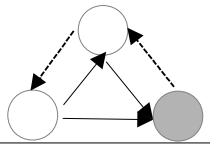
Figure 2.3: Within- and between-group densities by groups and waves

Table 2.2, shows descriptive statistics for tie creation and closure in dyads, two-paths and transitive triplet network structures for the three waves of our sample for different group compositions. Since there currently exists no tool to calculate these statistics, an R-function (see Appendix D) has been developed for this purpose. In collaboration with F. Agneessens, this is further developed in order to provide an openly accessible package that produces a dyad and triad census for in- and out-group ties by groups. For the counts of two-paths and triplets, the solid lines must exist, irrespective of whether the dashed (reciprocating) lines do or not. The columns Wave 1 to 3 show the cross-

sectional counts of realized structures, whereas the columns in between indicate the proportion of two-paths that have been closed between waves and thus forming a transitive triplet.[15] These presented network structures are differentiated for the intra- and inter-group case respectively, referring to the case where the two individuals that are connected by the respective network structure belong to the same or a different group (intra- and inter-group respectively).

With regard to dyads, we see that the average of existing and created ties within groups (intragroup) is higher than existing and created ties between groups (intergroup) throughout all waves. A similar pattern shows when looking at the presence of inter- and intra-ethnic two-paths and triplets. At all three observations, more intra- than inter-ethnic two-paths and triads are present. However, the descriptives of change over time - the proportion of transitively closed two-paths – show that inter-group two-paths close more often transitively. For the first period inter-group two-paths are more than twice as often transitively closed to triplets than intra-group two-paths, reducing to less than twice as often during the second time period.. This is surprising, since it means that pupils proportionally more often become friend of friends with someone from the ethnic out-group than the ethnic in-group.

15 This means that created ties transitively closing are explicitly not reciprocal. However looking at reciprocated two-paths and triads shows a similar pattern.

Dyads	Wave 1		Wave 2		Wave 3
Inter-group 		Creation		Creation	
	29.73	13.82	36.73	11.95	40.68
	(17.96)	(9.2)	(20.82)	(9.64)	(26.94)
Intra-group 					
	55.73	21.18	64.32	19.91	70.32
	(40.98)	(15.82)	(42.55)	(13.26)	(40.56)
Two paths and Triplets					
		Closure (in %)		Closure (in %)	
Intra-group 	203.28		289.91		325
	(142.18)		(196.03)		(225.16)
		3.07		2.53	
		(2.52)		(2.19)	
	59.32		98.36		108.41
	(60.69)		(78.83)		*(73.93*
Inter-group 	52.27		87.68		106
	(52.72)		(68.72)		(91.02)
		6.73		4.35	
		(7.22)		(5.69)	
	33.36		40.82		50.27
	(31.08)		(26.62)		(42.40)

*standard deviations in brackets

note: dotted lines can be, but must not be present

Table 2.2: Descriptives of intra- and inter-group dyads, twopaths and triplets

2.4.2 Hierarchical Stochastic Actor Orientented Models

As already mentioned a more accurate examination of homophily can be realized by applying dynamic models of network evolution. To empirically test the hypotheses, in table 2.3[16] we present estimates from a set of three basic models for 22 classrooms. The first model (**M1**) of table 2.3 can be seen as a baseline model to assess the extent of general homophily in dyads (**H1a**) and controlling for normal transitivity (**H2**) without accounting for specific group compositions in triads. The homophily effect models the contribution of both actors belonging to the same group (versus different groups) to the creation and maintenance of a tie and if found to be positive supports our first hypothesis (H1a). The transitivity parameter indicates the contribution of each intermediately connected actor to the formation and and maintenance of a tie, and if positive, supports our second main hypothesis (H2). In the second model (**M2**) homophily is differentiated for both groups by adding an interaction between the main effect indicating homophily and the main effect representing non-native egos. In this model, the homophily main effect indicates the tendency for native pupils to rather nominate other native pupils than non-natives. The linear combination of the non-native ego effect and it's interaction effect with the homophily parameter shows the homophily for the group of non-natives. The hypotheses that homophily for the group of native pupils is stronger (H1b) is confirmed if the aforementioned linear combination is negative. The third model (**M3**) additionally includes effects representing groupwise transitivity effects for both groups separately. Triadic constellations are modelled where ego *i* and alter *j* are from the same and respectively a different group, while the group association of the intermediate actor remains unspecified. Additionally, these effects for

16 This table only includes the core effects. Table A1 in the appendix gives a list of all effects (including controls) and the variances.

intra- and inter-group transitivity have been interacted with the variable indicating non-native ego's. In this specification, the main effect of intra-group transitivity indicates the contribution of intermediate actors for transitivity between two native actors (compared to the dyadic case). The linear combination of the interaction of intra-group transitivity and the indicator variable for non-native ego's thus represents the same for both actors being non-natives. If the interaction effect between non-native ego's and and intra-group transitivity is negative and significant, this supports the expectation that intra-group transitivity is larger for natives (H3a). Structural homophily for each group will be calculated for natives by subtracting the main effect of inter-group from intra-group transitivity. For transitivite homophily of non-natives one first has to add the respective interaction terms between the groupwise transitivity parameters and variable indicating non-native egos to them and then calculate the difference. After structural homophily for both groups has been calculated we subtract the value for non-natives from the value for natives. If this results in a positive significant value, we can not reject the expectation that natives show a stronger tendency for structural homophily (H3b).In the dyadic case, the effects for intra- and intergroup nominations are mutually exclusive; this is not the case for triadic constellations due to multiple possibilities of being connected via intermediate actors. Therefore both effects are present in the model and can be interpreted as the contribution of existing intermediate actors for the creation and maintenance of within- and between-group ties compared to the dyadic case.

		M1		M2		M3	
		b	s.e.	b	s.e.	b	s.e.
Homophily		.107	.025 **				
	Natives			.001	.040	.077	.064
	Non-natives			.156	.031 **	.179	.038 **
Transitivity		.660	.025 **	.654	.052 **		
	natives intra-group					.637	.024 **
	Non-natives intra-group					.668	.025 **
	native inter-group					.679	.023 **
	Non-native inter-group					.664	.026 **

**Distribution of parameters does not co * Less then .05 of distribution covers 0

+ Table includes only core effects; full model specification can be found in appendix Table A2

Table 2.3: Linear combinations of effects from hierarchichal SAOM +

The baseline model (**M1**) provides support for general choice homophily (**H1a**), with the respective parameter being positive (b=0.107, se=0.025), as well for a general tendency of transitivity (b=0.660, se=0.025) which supports **H2**. By calculating the odds-ratios we can see that pupils are 1.1 times more likely to nominate an alter from the same ethnic group and 1.94 times as likely to nominate a friend that they are indirectly connected to. Differentiating for group specific homophily, the second model (**M2**) shows a positive significant homophily effect for non-native pupils (b=0.156, se=0.031), while the effect for natives proves not to be significant (b= 0.001, p=0.040). The conditional odds ratio of the homophily effect for non-natives reveals that they have a 1.17 times higher chance to nominate an in-group versus an out-group friend. This supports the hypothesis of group differences with regard to dyadic homophily, however in opposite direction as expected in **H1b**. It can be argued that the general homophily parameter from the previous has been driven by the homophily of non-natives and is lower because of the non-significant effect for natives.[17]

17 Since HSAOM is in the family of generalized linear models, this increase might be due to unobserved heterogeneity (see Mood, 2010).

In the last model (**M3**), parameters for intra- and intergroup transitivity have been added and interacted with the indicator for non-native pupils. These effects indicate whether there is a difference for non-natives with regard to the respective transitive composition compared to natives.[18] Again for a more intuitive understanding of the parameters, we present the linear combination of these, showing natives and non-natives actual tendency for transitivity. Looking at the results of **M3** a tendency of intra-group transitivity for both groups(natives: b=0.637, se=0.24; p<.001; non-natives: b=0.668, se=0.25, p<.001). The results for the interaction effect indicating non-native ego's and intra-group transitivity which shows the difference for non-native's intra-group transitivity compared to natives is positive and indicatates a significant difference (see Appendix table A1[19]: b= 0.031, se=0.009, p<.001). While natives have 1.89 higher chance to have an intragroup tie when it is embedded in a triad, it is 1.95 times higher for non-natives.[20] While this supports the idea that both groups differ, it shows to be in the opposite direction as expected and thus contradicts our hypothesis (**H3a**). Regarding intergroup transitivity, we do not find a significant difference between both groups (see Appendix table A1[21]: b= -0.015, se=0.013). Lastly, in order to compare structural homophily between both groups we first calculate the difference of ingroup transitivity

18 For the dyadic case, the resulting model includes an effect representing choice homophily for natives (*sameX*) an effect for ego being non-native (*egoX*) and an interaction of those which represents non-natives tendency to befriend other non-natives (*egoXsameX*) Consequently from the 2 by 2 possible combinations of dyads between native and non-native ego's and alters, this leaves the dyad of native ego and non-native alter as reference category. To present more intuitive measures of dyadic homophily, the respective linear combinations of these effects are calculated, so that table 2.3 reports the dyadic homophily parameter for each group. The linear combination of non-native homophily results from sameX_migrant and the egoXsameX_migrant interaction. This is similar for the triadic constellations, where the reported effects are linear combinations showing each groups tendency for in-group and out-group transitivity. Non-natives´ intragroup transitivity results from the combination of the sameXtransTrip_migrant and the egoXsameXtranstrip_migrant (diffX respectively for the intergroup case). The raw effect sizes of all parameters estimated in the models can be found in table A1 in the appendix.

19 Effect name: egoXsameXTransTrip_migrant

20 However, this difference is very small and because the estimated effects are fixed, the standard error might be underestimated.

21 Effect name: egoXdiffXTransTrip_migrant

for natives and non-natives seperately as outlined earlier. For natives, taking the difference between intragroup transitivity (b1=0.637, se=0.024 and b2=.679, se=.023) shows a small but significant negative effect (b_{nat}= -.048, se=.016, p < .001) indicating that natives show a small tendency for inter-group over intra-group transitivity (*structural heterophily*). Calculating the differences for non-natives intra-group (b3=0.668, se=0.25) and inter-group transitivity (b4=.664, se=.026) results in a not-significant small positive effect ($b_{non\text{-}nat}$=.004, sd=.009, p=.661). Calculating the difference of structural homophily for natives and non-natives results in a significant negative effect (b_{diff}= -.052, se=.017, p=.001). This speaks against our expectation that natives have a stronger tendency for structural homophily (H3b), but actually refers to natives' rather closing transitively to non-natives than to other natives, while non-natives show no difference in transitivity based on the group of the respective alter.

This difference in effect sizes might appear small, especially in comparison to the effect size of the respective transitivity parameters. However, one has to consider that the transitivity parameter applied here models a linear contribution of each intermediate actor. Thus, it assumes that for each intermediately shared friend the respective transitivity becomes more attractive. Regarding structural heterophily of natives, one can consider the example of a non-native pupil having some ties to a group or clique of strongly connected natives. If this clique consists of five native pupils who all are friends with each other, four of them already being and the fifth becoming friends with the non-native peer this would provide four intermediate actors contributing to this transitive closure. By multiplication this results in a reduction of the log odds-ratio by about a third (b_{nat}(-.048 * 4) = -.232), translating to this native pupils' odds of choosing a friendship with another native peer who is similarly embedded being lower by a factor of .8l than choosing a friendship with the non-native alter. Given that the non-native

group does not show this tendency for structural heterophily, this finding indicates a small but not necessarily irrelevant difference between groups.

2.5 Conclusions

The first part of our analysis supports dyadic choice homophily when not differentiating for natives and non-natives and finds that transitivity contributes to tie creation and maintenance. Differentiating tendencies for homophily and groupwise transitivity between the group of native and non-native pupils demonstrates that dyadic homophily differed between both groups, though surprisingly only the group of non-natives demonstrate a preference for in-group ties in dyads. Lastly, we find some evidence for group differences of structural homophily between both groups. To be more precise, natives showed a small but significant tendency for transitive heterophily, indicating that natives rather have embedded friendships to the out-group than to their own. Non-natives on the other hand, did not show any difference for transitivity with respect to the alter being from their own or the out-group.

One major contribution of this study is the longitudinal investigation of processes leading to ethnic friendship segregation in classrooms. We argue that a cross-sectional perspective is only able to observe the extent of segregation at a specific time point, while one of it's possible causes, namely homophily, is a mechanism that is observed over time. However, by analysing the evolution of networks over time, we are not only able to model the actual process of homophily as a mechanism of rather creating and maintaining ties to the own than to the out-group, but at the same time we can shed light on network inherent dynamic processes. One of these processes is transitivity, pupils tendency to become friends with the friends of their friends. TThis may be interrelated with homophily and also contribute to segregation in networks. Until now only a few studies exist which try to disentangle both mechanisms. One example poses the study of

Grund and Densley (2015), who apply as a cross-sectional approach. In line with their findings, we also observe that transitive triplets between pupils of the same background are more frequent than transitive triplets connecting pupils of different background. However, looking at changes pupils make in their network, descriptives and results from our model estimations did not show evidence for intra-group two-paths to have a higher tendency for closure than those connecting pupils of different background. In fact, we even find that, natives transitively close two-paths proportionally more often to non-natives rather than to their own. Intuitively, finding transitive triplets to be more present within groups but not finding a tendency for in-group transitive closure might seem to contradict each other. However, it outlines the difference between analysing the state of a network and a process shaping it.

This difference and the role of endogenous network processes for it can be better understood when interpreteding our result in line with the arguments made for the emergence of segregation by Grund (2014). He points out that networks segregate partly because of actors' in-group preference for tie creation and partly due to in-group transitive closure. When the first ties of a (new) network emerge and actors only have a slight preference for their in-group, they will become more likely to be indirectly connected to in-group others than to those from the out-group. Given the well-established tendency for friends of friends to become friends, this can further amplify initial homophily to lead to even more segregation in observed networks. This is because ties established due to transitive closure are more frequent to other in-group members. This is even the case when proportionally there is no stronger tendency to close triads with in-group than with out-group members. In other words, friends of friends become friends, whether they are from the same or a different ethnic group. Still, this does not necessarily prevent networks from segregating because a slight in-

group preference at early stages of a network provides numerically more in-group two-paths. Even if in-group two-paths do not proportionally close more often (not showing structural homophily), numerically more in-group ties can be created.

Given that the first wave of our data was collected about half a year after the classrooms had been assembled we observed that the formed friendships were already considerably segregated. This can reflect that the above mentioned process of ethnic friendship segregation in the emergence of the here observed networks has already progressed to some extent after only half a year. However, the descriptives indicated a stronger network growth, in terms of increasing average out-degree, for the first period while the second period did not show this to such an extent. For our data, this could mean that pupils which are assembled into a classroom half a year before data collection started to mainly create friendships with each other. This continued to a point in time after the first and before the second wave of data collection, where the increase of pupils' average friendship nominations reduced to smaller differences. This can suggest that initially pupils make additional friends until their need for social relations is saturated to some extent. In this phase of network growth, processes of segregation in emerging networks such as described by Grund (2014) can occur. However, for the second period between observations, the networks show only little growth. This might indicate that processes of tie creation have less impact on the network evolution than process of tie maintenance have. After pupils made a certain amount of friends, for example based on homophily of salient characteristics such as migration background, they might learn more about their friends less-salient characteristics such as attitudes or preferences. If differences in these less visible characteristics become apparent, they might decide against maintaining these friendships. In that respect, our statistical model relies on strict assumptions which do not reflect such time heterogeneity in network effects. Neither do we differentiate

mechanisms of friendship creation and maintenance, nor do we test for our modelled parameters to differ between both periods. While this is technically possible within the SAOM framework, it was beyond the aim of this study which seeks to investigate in-group tendencies of transitive closure.

However, one of the contributions of this work was to relax the assumption that homophily is similar for the group of natives and non-natives. When differentiating homophily between the two groups we revealed the unexpected finding that general homophily is driven by non-natives to have higher chances of nominating friends from the in-group than the out-group, whereas natives don't show such a tendency. This phenomenon might be better understood relating to the findings in figure 2.3. where it was shown that at the first wave non-natives proportionally had more out-group than in-group nominations, which was not the case for the group of natives. This might be due to differences in group wise reciprocity. Meaning that non-natives initially nominated more natives than other non-natives, but that these between-group ties have not been reciprocated from natives. This could even be understood as natives rejecting attempts of non-natives to become friends. The observation that at the second wave non-natives now as well show proportionally more in-group friendships could be interpreted as a reaction to the rejection of their efforts. Consequently, non-natives dissolve the non-reciprocated ties and instead form more ties to other non-natives, leading to our finding of a homophily effect that is driven by the group of non-natives. If this is the case, it implies that the process of creating group boundaries (see Wimmer, 2013) is initiated from the group of natives - potentially due to negative out-group attitudes, prejudice and different forms of social distance - after which, non-natives perceive these boundaries and reinforce them as well. A potential explanation for this finding might be deduced from self-categorisation theory (Turner et al., 1987). It argues that categorization is

strongly dependent on the context and situation. In line with the theoretical considerations from which our hypothesis about group differences in homophily is derived, homophily is not driven by natives group identification as natives. In contrast, the pattern of change that we observe in our network results from non-natives starting to identify as such. While at first observation non-natives do not self-define as a common non-native group, natives differentiate themselves as one common group thus consolidating, accentuating and signalling the group boundary to the others right from the start. This might then lead non-natives to have solid grounds to adopt this categorization and likewise set themselves apart from the natives subsequently. Therefore, they develop a tendency to prefer tie creation and maintenance with their newly defined in-group of non-natives. This means that the group of natives which in our sample constitutes the majority of pupils in most of the classrooms, establishes boundaries towards the group of non-natives. These boundaries could be reflected in day to day interactions, ranging between rather spending time on the playground with their own than with non-natives pupils to for example directly rejecting an invitation to a birthday celebrations from a non-native classmate. While non-native pupils initially seemingly are motivated to integrate themselves by forming bonds with their native peers, their efforts are one-sided and considered to not pay off.

On a first glance, this interpretation does not go in line with the finding of structural heterophily, non-natives tendency to rather be transitively connected to non-natives than to natives. However, it can be argued that finding this tendency is driven by differences in the composition of classrooms, namely those where only very few non-native pupils are present. Here, it could be the case that that a clique of interconnected natives connects to one of the few non-native pupils, providing several intermediate actors that contribute to the modelled parameter of natives' inter-group transitivity while providing

less counts to the larger proportion of other natives they are not friends with. Such a scenario could be interpreted in line with Feld's (1981) theory of social foci. He argues that pupils who share several social *foci* such as sport clubs or religious groups, assess the relation with each other as more positive thus being more likely to consider each other as friends. However, the propensity of two pupils to share such *foci* can be subject to housing segregation. Homophily might be a reflection of ethnically segregated housing because pupils of native background are more likely to live in the same area, therefore share more social *foci* and thus are more likely to be friends with another. Finding natives tendency for structural heterophily might be the result of a few non-native pupils which live in a native majority area and share one or several *foci* with a larger clique of natives from their neighbourhood. However,

Even though our study presents findings that provide insight into the possible processes of friendship segregation, the question arises about how generalizable the conclusions are. First of all, this study only examines one specific type of personal relation. Windzio and Bicer (2013) make the point that for friendship networks, the tendencies for segregation are lowest compared to other more costly relations (e.g. leisure time activities, visits at home, and contact between the parents). With regard to the generalizability of our findings, it would be of interest to explore how choice homophily and transitive closure relate and contribute to the evolvement of segregation in networks of more risk-inherent relationships. When reproducing our work, higher segregation in more costly network ties could result from the combinations of following extremes: firstly, a stronger initial tendency for dyadic homophily when the first ties are created, with no in-group preference for triadic closure: secondly, lower or no initial tendency for dyadic homophily, but a strong preference for in-group others when closing triads transitively: thirdly, in an extreme case of both processes operating, a strong initial

dyadic homophily and an in-group preference for transitive triadic closure. Moreover, limited generalizability of our results can also be due to sample restrictions. Because the data stems only from the area of two German city, our results might be driven by regional particularities. However, most prior longitudinal network studies looked at smaller samples than our case of 22 networks. We therefore argue that our findings are less likely to result from only a few particular networks.

Concluding, we show that the differentiation of choice homophily between native and non-native pupils can provide valuable insights about the process of network segregation. Showing that in the classrooms we investigated, homophily is driven by the non-native group, underscores the importance of taking group differences into account when homophily and segregation is researched in network studies. Furthermore, not finding a tendency for structural homophily in the presence of dyadic homophily generally supports our theoretical idea of intergroup uncertainties to be attenuated through an intermediate actor. Future research relating to our work can provide additional insights especially in the following ways:

- by collecting individual data about real group related uncertainties and attitudes, and by investigating how these are related to the evolvement of peoples´ relational networks, in order to test the implications of our theoretical model.
- by looking at how mechanisms of choice and structural homophily contribute for segregation in other more risk-laden relational network types (see Windzio and Bicer, 2013).
- by looking at different regional contexts, e.g. for other countries with a different history of migration and with other ethnic majority and minority groups.

Another line of research could provide insights by focussing on patterns of group-wise

reciprocity and tie dissolution which was beyond the scope of this study. In addition, by investigating a larger and more representative sample of networks, applying the multilevel methods for analysing longitudinal data (such as HSAOM) can allow to relax assumptions made in the modelling choices applied here by investigating time and context homogeneity of network mechanisms.

Ideally these questions should be linked to a systematic analyses of potentially relevant contextual factors (e.g. ethnic composition of classrooms, level of segregation, teaching styles or other relevant network or context-level characteristics). This work underlines the importance of taking differences into account, when investigating questions of homophily and segregation and we encourage other research to diversify and extend their contents of analyses to other network data and contexts.

2.6 Literature

Agneessens, F.; Skvoretz, J. (2012): Group differences in reciprocity, multiplexity and exchange: measures and application. In: *Quality & Quantity* 46 (5), 1523–1545.

Ayres, J. (1979): Uncertainty and social penetration theory expectations about relationship communication: A comparative test. In: *Western Journal of Speech Communication* 43 (3), 192–200.

Bade, Klaus J.; Oltmer, Jochen (2010): Deutschland. In: Klaus J. Bade, Pieter C. Emmer, Leo Lucasen, and Jochen Oltmer (Eds.): Enzyklopädie Migration in Europa: vom 17. Jahrhundert bis zur Gegenwart. Paderborn: Schöningh, pp. 141–170.

Baerveldt, Chris; van Duijn, Marijtje A. J.; Vermeij, Lotte; van Hemert, Dianne A. (2004): Ethnic boundaries and personal choice. Assessing the influence of individual inclinations to choose intra-ethnic relationships on pupils´ networks. In: *Social Networks* 26, 55–74.

Baerveldt, Chris; Völker, Beate; van Rossem, Ronan (2008): Revisiting Selection and Influence: An Inquiry into the Friendship Networks of High School Students and Their Association with Delinquency. In: *Canadian Journal of Criminology and Criminal Justice* 50 (5), 559–587.

Barrera, Davide (2005): Trust in Embedded Settings. Proefschrift. Universiteit Utrecht, Utrecht.

Barrera, Davide; Buskens, Vincent (2009): Third-Party Effects. In: Karen S. Cook,

Chris Snijders, Vincent Buskens and Coye Cheshire (Eds.): eTrust: Forming Relationships in the Online World: Russell Sage Foundation, pp. 37–72.

Baumeister, Roy F.; Leary, Mark R. (1995): The Need to Belong: Desire for Interpersonal Attachments as a Fundamental Human Motivation. In: *Psychological Bulletin* 117 (3), 497–529.

Berger, Charles R.; Calabrese, Richard J. (1975): Some Explorations in Initial Interaction and Beyond: Toward a Developmental Theory of Interpersonal Communication. In: *Human Communication Research* 1 (2), 99–112.

Berger, Charles R.; Bradac, James J. (1982): Language and social knowledge. Uncertainty in interpersonal relations. London: Arnold.

Berger, Charles R.; Kellermann, Kathy (1994): Acquiring social information. In: John A. Daly and John M. Wiemann (Eds.): Strategic Interpersonal Communication. Hillsdale, NJ: Lawrence Erlbaum, pp. 1–32.

Bettencourt, Ann B.; Dorr, Nancy; Charlton, Kelly; Hume, Deborah L. (2001): Status differences and intergroup bias: A meta-analytic examination of the effects of status stability, status legitimacy and group permeability. In: *Psychological Bulletin* 127, 520–542.

Block, Per (2015): Reciprocity, transitivity, and the mysterious three-cycle. In: *Social Networks* 40, 163–173.

Block, Per; Grund, Thomas (2014): Multidimensional homophily in friendship networks. In: *Network Science* 2 (2), 189–212.

Borgatti, Stephen P.; Everett, Martin G.; Johnson, Jeffrey C. (2013): Analyzing Social Networks. London: Sage Publications Ltd.

Buchan, Nancy R.; Croson, Rachel T. A.; Dawes, Robyn M. (2002): Swift Neighbors and Persistent Strangers: A Cross-Cultural Investigation of Trust and Reciprocity in Social Exchange. In: *American Journal of Sociology* 108 (1), 168–206.

Burt, Ronald S.; Knez, Marc (1995): Kinds of Third-Party Effects on Trust. In: *Rationality and Society* 7 (3), 255–292.

Buskens, Vincent (2002): Social Networks and Trust. Boston, Mass.: Kluwer Academic Publishers.

Cartwright, Dorwin; Harary, Frank (1956): Structural balance. A generalization of Heider's theory. In: *Psychological Review* 63 (5), 277–293.

Çelebi, Elif; Verkuyten, Maykel; Köse, Talha; Maliepaard, Mieke (2014): Out-group trust and conflict understandings: The perspective of Turks and Kurds in Turkey. In: *International Journal of Intercultural Relations* 40, 64–75.

Coleman, James S. (1990): Foundations of social theory. Cambridge, Mass.: Belknap Press of Harvard Univ. Press.

Darley, John M.; Gross, Paget H. (1983): A hypothesis-confirming bias in labeling

effects. In: *Journal of Personality and Social Psychology* 44 (1), 20–33.

Davis, James A. (1967): Clustering and Structural Balance in Graphs. In: *Human Relations* 20 (2), 181–187.

Davis, James A.; Leinhardt, Samuel (1972): The structure of positive interpersonal relations in small groups. In: Joseph Berger (Ed.): Sociological theories in progress. Boston, Mass.: Houghton Mifflin (Volume 2), pp. 218–251.

Denissen, Jaap J. A.; Penke, Lars; Schmitt, David P.; van Aken, Marcel A. G. (2008): Self-Esteem Reactions to Social Interactions: Evidence for Sociometer Mechanisms Across Days, People, and Nations. In: *Journal of Personality and Social Psychology* 95 (1), 181–196.

Derlega, Valerian J.; Winstead, Barbara A.; Greene, Kathryn (2008): Self-Disclosure and Starting a Close Relationship. In: Susan Sprecher, Amy Wenzel und John Harvey (Eds.): Handbook of relationship initiation. Hove, UK: Psychology Press, pp. 153–174.

Dhont, Kristof; van Hiel, Alain (2011): Direct contact and authoritarianism as moderators between extended contact and reduced prejudice: Lower threat and greater trust as mediators. In: *Group Processes & Intergroup Relations* 14 (2), 223–237.

Dijkstra, Jan Kornelis; Berger, Christian; Lindenberg, Siegwart (2011): Do physical and relational aggression explain adolescents' friendship selection? The competing roles of network characteristics, gender, and social status. In: *Aggressive Behavior* 37 (5), 417–429.

Duck, S.; Miell, D. (1986): Charting the development of personal relationships. In: R. Gilmour and S. Duck (Eds.): The emerging field of personal relationships. Hillsdale, NJ: Lawrence Erlbaum, pp. 133–143.

Fehr, Beverley (2008): Friendship Formation. In: Susan Sprecher, Amy Wenzel and John Harvey (Eds.): Handbook of relationship initiation. Hove, UK: Psychology Press, pp. 29–54.

Fein, Steven; Spencer, Steven J. (1997): Prejudice as Self-Image Maintenance: Affirming the Self through Derogating Others. In: *Journal of Personality and Social Psychology* 73 (1), 31–44.

Feld, S. (1981). The Focused Organization of Social Ties. *American Journal of Sociology, 86*(5), 1015-35

Friemel, Thomas N. (2012): Network dynamics of television use in school classes. In: *Dynamics of Social Networks (2)* 34 (3), 346–358.

Furman, Wyndol; Buhrmester, Duane (1992): Age and Sex Differences in Perceptions of Networks of Personal Relationships. In: *Child Development* 63 (1), 103–115.

Goodreau, Steven M; Kitts, James A.; Morris, Martina (2009): Birds of a Feather, or Friend of a Friend? Using Exponential Random Graph Models to Investigate Adolescent Social Networks. In: *Demography* 46 (1), 103–125.

Gould, Roger V.; Fernandez, Roberto M. (1989): Structures of Mediation: A Formal Approach to Brokerage in Transaction Networks. In: *Sociological Methodology* 19, 89–126.

Grund, Thomas U. (2014) Network Size and Homophily: Same-Sex Friendships in 595 Scandinavian Schools, in Manzo, G. (eds.) Analytical Sociology. Norms, Actions, Networks. New York: Wiley, pp. 299-316.

Grund, Thomas U.; Densley, James A. (2015): Ethnic Homophily and Triad Closure: Mapping Internal Gang Structure Using Exponential Random Graph Models. In: *Journal of Contemporary Criminal Justice* 31 (3), 354-370.

Hallinan, Maureen T.; Williams, Richard A. (1989): Interracial Friendship Choices in Secondary Schools. In: *American Sociological Review* 54 (1), 67–78.

Hammer, Muriel (1979): Predictability of social connections over time. In: *Social Networks* 2 (2), 165–180.

Hartup, Willard W.; Stevens, Nan (1997): Friendships and Adaptation in the Life Course. In: *Psychological Bulletin* 121 (3), 355–370.

Heider, Fritz (1946): Attitudes and cognitive organization. In: *Journal of Psychology* 21, 107–112.

Helsen, Marianne; Vollebergh, Wilma; Meeus, Wim (2000): Social Support from Parents and Friends and Emotional Problems in Adolescence. In: *Journal of Youth and Adolescence* 29 (3), 319-335.

Holland, Paul W.; Leinhardt, Samuel (1970): A Method for Detecting Structure in Sociometric Data. In: *American Journal of Sociology* 76 (3), 492–513.

Holland, Paul W.; Leinhardt, Samuel (1971): Transitivity in Structural Models of Small Groups. In: *Comparative Group Studies* 2 (2), 107–124.

Insko, Chester A.; Schopler, John (1998): Differential Distrust of Groups and Individuals. In: Constantine Sedikides, John Schopler and Chester A. Insko (Eds.): Intergroup Cognition and Intergroup Behavior. Mahwah, N.J: Lawrence Erlbaum Associates, pp. 75–107.

Koskinen, Johan H.; Snijders, Tom A. B. (2007): Bayesian inference for dynamic social network data. In: *Journal of Statistical Planning and Inference* 137 (12), 3930–3938.

Koskinen, Johan H.; Snijders, Tom A. B. (2019): Multilevel longitudinal analysis of social networks (in preparation).

Kossinets, Gueorgi; Watts, Duncan J. (2009): Origins of Homophily in an Evolving Social Network. In: *American Journal of Sociology* 115 (2), 405–450.

Krackhardt, David; Stern, Robert N. (1988): Informal Networks and Organizational Crises: An Experimental Simulation. In: *Social Psychology Quarterly* 51 (2), 123–140.

Lazarsfeld, Paul F.; Merton, Robert K. (1954): Friendship as a Social Process: A Substantive and Methodological Analysis. In: Morroe Berger, Theodore Abel and Charles H. Page (Eds.): Freedom and Control in Modern Society. Toronto: Van Nostrand, pp. 18–66.

Levesque, Maurice J.; Steciuk, Mark; Ledley, Charles (2002): Self-Disclosure Patterns among Well-Acquainted Individuals: Disclosers, Confidants and Unique Relationships. In: *Social Behavior and Personality* 30 (6), 579–592.

Lewis, Kevin (2015): How Networks Form: Homophily, Opportunity, and Balance. In: Robert A. Scott und Stephen M. Kosslyn (Eds.): Emerging Trends in the Social and Behavioral Sciences: John Wiley & Sons, Inc, pp. 1–14.

Lubbers, Miranda J.; Molina, José Luis; Lerner, Jürgen; Brandes, Ulrik; Ávila, Javier; McCarty, Christopher (2010): Longitudinal analysis of personal networks. The case of Argentinean migrants in Spain. In: *Dynamics of Social Networks* 32 (1), 91–104.

Lubbers, Miranda J.; Snijders, Tom A. B.; van der Werf, Margaretha P. C. (2011): Dynamics of Peer Relationships Across the First Two Years of Junior High as a Function of Gender and Changes in Classroom Composition. In: *Journal of Research on Adolescence* 21 (2), 488–504.

Macy, Michael W.; Skvoretz, John (1998): The Evolution of Trust and Cooperation between Strangers: A Computational Model. In: *American Sociological Review* 63 (5), 638–660.

Marsden, Peter V. (1982): Brokerage Behavior in Restricted Exchange Networks. In: Peter V. Marsden and Nan Lin (Eds.): Social structure and network analysis. 2. Edition, Beverly Hills, Calif: Sage, pp. 201–218.

McPherson, Miller; Smith-Lovin, Lynn (1987): Homophily in Voluntary Organizations: Status Distance and the Composition of Face-to-Face Groups. In: *American Sociological Review* 52 (3), 370–379.

McPherson, Miller; Smith-Lovin, Lynn; Cook, James M. (2001): Birds of a Feather: Homophily in Social Networks. In: *Annual Review of Sociology* 27, 415–444.

Meeus, Wim; Dekovic, Maja (1995): Identity Development, Parental and Peer Support in Adolescence: Results of a National Dutch Survey. In: *Adolescence* 30 (120), 931-944.

Miell, Dorothy; Duck, Steve (1986): Strategies in Developing Friendships. In: Valerian J. Derlega and Barbara A. Winstead (Eds.): Friendship and social interaction. New York: Springer (Springer series in social psychology), pp. 129–143.

Mood, Carina (2010): Logistic Regression: Why We Cannot Do What We Think We Can Do, and What We Can Do About It. In: *European Sociological Review* 26 (1), 67–82.

Moody, James (2001): Race, School Integration, and Friendship Segregation in America. In: *The American Journal of Sociology* 107 (3), 679–716.

Moody, James (2009): Network dynamics. In: Peter Hedström and Peter S. Bearman

(Eds.): The Oxford handbook of analytical sociology. Oxford: Oxford Univ. Press, pp. 447–474.

Munniksma, Anke; Stark, Tobias H.; Verkuyten, Maykel; Flache, Andreas; Veenstra, René (2013): Extended intergroup friendships within social settings: The moderating role of initial out-group attitudes. In: *Group Processes & Intergroup Relations* 16 (6), 752–770.

Munniksma, Anke; Verkuyten, Maykel; Flache, Andreas; Stark, Tobias H.; Veenstra, René (2015): Friendships and out-group attitudes among ethnic minority youth: The mediating role of ethnic and host society identification. In: *International Journal of Intercultural Relations* 44, 88–99.

Pinter, Brad; Greenwald, Anthony G. (2011): A comparison of minimal group induction procedures. In: *Group Processes & Intergroup Relations* 14 (1), 81–98.

Prisbell, Marshall; Andersen, Janis F. (1980): The importance of perceived homophily, level of uncertainty, feeling good, safety, and self-disclosure in interpersonal relationships. In: *Communication Quarterly* 28 (3), 22–33.

Ruth M. Ripley, Tom A.B. Snijders, Zsófia Boda, Andras Vörös, and Paulina Preciado (2018). Manual for SIENA version 4.0. Oxford: University of Oxford, Department of Statistics; Nuffield College.

Rivera, Mark T.; Soderstrom, Sara B.; Uzzi, Brian (2010): Dynamics of Dyads in Social Networks: Assortative, Relational, and Proximity Mechanisms. In: *Annual Review of Sociology* 36, 91–115.

Scheepers, Daan; Spears, Russell; Doosje, Bertjan; Manstead, Antony S. R. (2006): The social functions of in-group bias: Creating, confirming, or changing social reality. In: *European Review of Social Psychology* 17 (1), 359–396.

Schneider, Silke L. (2008): Anti-Immigrant Attitudes in Europe: out-group Size and Perceived Ethnic Threat. In: *European Sociological Review* 24 (1), 53–67.

Semyonov, Moshe; Raijman, Rebeca; Gorodzeisky, Anastasia (2006): The Rise of Anti-Foreigner Sentiment in European Societies, 1988-2000. In: *American Sociological Review* 71 (3), 426–449.

Smith, Sanne; Maas, Ineke; van Tubergen, Frank (2014): Ethnic in-group friendships in schools: Testing the by-product hypothesis in England, Germany, the Netherlands and Sweden. In: *Social Networks* 39, 33–45.

Snijders, Tom A. B. (2001): The Statistical Evaluation of Social Network Dynamics. In: *Sociological Methodology*, 361–395.

Snijders, Tom A. B.; Koskinen, Johan; Schweinberger, Michael (2010a): Maximum likelihood estimation for social network dynamics. In: *The Annals of Applied Statistics* 4 (2), 567–588.

Snijders, Tom A. B.; van Bunt, Gerhard G de; Steglich, Christian E. G. (2010b): Introduction to Stochastic Actor-Based Models for Network Dynamics. In: *Social Networks* 32, 44–60.

Sprecher, Susan; Felmlee, Diane (2008): Insider Perspectives on Attraction. In: Susan Sprecher, Amy Wenzel and John Harvey (Eds.): Handbook of relationship initiation. Hove, UK: Psychology Press, pp. 297–313.

Stark, Tobias H. (2015): Understanding the Selection Bias: Social Network Processes and the Effect of Prejudice on the Avoidance of out-group Friends. In: *Social Psychology Quarterly* 78 (2), 127–150.

Stark, Tobias H.; Flache, Andreas; Veenstra, René (2013): Generalization of Positive and Negative Attitudes Toward Individuals to out-group Attitudes. In: *Personality and Social Psychology Bulletin* 39 (5), 608–622.

Tajfel, Henri (1970): Experiments in Intergroup Discrimination. In: *Scientific American* 223, 96–102.

Tajfel, Henri; Turner, John C. (1986): 1. The Social Identity Theory of Intergroup Behaviour. In: S. Worchel and W. G. Austin (Eds.): Psychology of intergroup relations. 2. Edition. Chicago: Nelson-Hall, pp. 7–24.

Tam, Tania; Hewstone, Miles; Kenworthy, Jared; Cairns, Ed (2009): Intergroup Trust in Northern Ireland. In: *Personality and Social Psychology Bulletin* 35 (1), 45–59.

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

Uehara, Edwina (1990): Dual Exchange Theory, Social Networks, and Informal Social Support. In: *American Journal of Sociology* 96 (3), 521–557.

Verbrugge, Lois M. (1977): The Structure of Adult Friendship Choices. In: *Social Forces* 56 (2), 576–597.

Vermeij, Lotte; van Duijn, Marijtje A. J.; Baerveldt, Chris (2009): Ethnic segregation in context: Social discrimination among native Dutch pupils and their ethnic minority classmates. In: *Social Networks* 31, 230–239.

Wagner, Ulrich; Christ, Oliver; Pettigrew, Thomas F. (2008): Prejudice and Group-Related Behavior in Germany. In: *Journal of Social Issues* 64 (2), 403–416.

Weeks, Trisha L.; Pasupathi, Monisha (2010): Autonomy, Identity, and Narrative Construction with Parents and Friends. In: Kate C. McLean and Monisha Pasupathi (Eds.): Narrative Development in Adolescence. Creating the Storied Self. New York: Springer Science+Business Media (Advancing Responsible Adolescent Development), pp. 65–91.

Wimmer, Andreas (2013): Ethnic boundary making. Institutions, power, networks. Oxford: Univ. Press (Oxford studies in culture and politics).

Wimmer, Andreas; Lewis, Kevin (2010): Beyond and Below Racial Homophily: ERG Models of a Friendship Network Documented on Facebook. In: *American Journal of Sociology* 116 (2), 583–642.

Windzio, Michael; Bicer, Enis (2013): Are we just friends? Immigrant integration into

high- and low-cost social networks. In: *Rationality and Society* 25 (2), 123–145.

Wright, Stephen C.; Aron, Arthur; McLaughlin-Volpe, Tracy; Ropp, Stacy A. (1997): The extended contact effect: Knowledge of cross-group friendships and prejudice. In: *Journal of Personality and Social Psychology* 73 (1), 73–90.

Yamagishi, Toshio; Kiyonari, Toko (2000): The Group as the Container of Generalized Reciprocity. In: *Social Psychology Quarterly* 63 (2), 116–132.

Yuki, Masaki; Maddux, William W.; Brewer, Marilynn B.; Takemura, Kosuke (2005): Cross-Cultural Differences in Relationship- and Group-Based Trust. In: *Personality and Social Psychology Bulletin* 31 (1), 48–62.

Yule, G. Udny (1912): On the Methods of Measuring Association Between Two Attributes. In: *Journal of the Royal Statistical Society* 75 (6), 579–652.

3 WHO BECOMES AND WHO STAYS FRIENDS? RELATING ETHNIC HOMOPHILY AND RECIPROCITY IN FRIENDSHIP CREATION AND MAINTENANCE IN CLASSROOMS

Co-Authors:

Prof. Jan O. Jonsson, Nuffield College, University of Oxford
Email: janne.jonsson@nuffield.ox.ac.uk

Dr. Isabel Raabe, Institute of Sociology, University of Zurich
Email: isabel.raabe@uzh.ch

3.1 Introduction

Many Western countries nowadays are high-immigration societies, and the integration of young people with ethnic minority background is often seen as a challenge, attracting considerable political and research interest. While much research has been targeting so-called structural integration of youth in the form of educational attainment (e.g., Heath and Brinbaum 2014; Heath et al. 2008), or cultural integration in terms of host-country identity (Koopmans 2005; Shore 2000), research on social integration defined in terms of interethnic friendships and relations has more recently received attention (e.g. Munniksma et al. 2013; Stark 2015, Leszczensky et al. 2015, Boda and Neray 2015). A lack of such integration may nevertheless lead to ethnic majority and minority populations living 'parallel lives', a lack of understanding between ethnic groups, and persistent prejudice and discrimination (e.g., Pettigrew and Tropp 2006). Ethnic minorities may as a consequence find it difficult to get ahead socioeconomically, as they lack the necessary social connections (Granovetter 1973; Lin 1990). Understanding what facilitates or complicates interethnic relationships is thus of great value for issues of immigrant integration and social stratification, and potentially also for social

cohesion. While interethnic friendships are valuable in their potential to reduce prejudice and provide social capital, they are circumscribed by two general phenomena, ethnic segregation (residential and in school) as well as ethnic homophily, originating from a general tendency for individuals to form relationships with people who are similar to themselves (McPherson et al., 2001). Thus, the commonly found ethnic friendship segregation stems from the limited chances to meet people of different ethnicity, and given those structural opportunities also from the propensity of sticking to one's own ethnic group – which is in itself a mixture of preference and external restrictions such as prejudice and discrimination.

Moreover, beneficial effects (such as social capital) of realised interethnic social relations are limited if they are short-lived and unstable (Rude and Herda 2010). Contact theory implicitly assumes that friendships are sustained over time and social capital as provided by an interethnic friendship will also be much more effective when built on a long-lasting friendship. This is especially the case as each friend is likely to bring about a number of acquaintences, allowing enduring friendships to gradually grow into wider social networks including many 'weak ties' (Granovetter 1973). However, there is only limited previous research on the stability of interethnic friendships. To our knowledge, only three papers have investigated this, and all of them have done so for the context of the racial segregation in the United States (Hallinan and Williams 1987; Aboud, Mendelson and Purdy 2003; Rude and Herda 2010). These studies all find that inter-racial friendships are less stable than intra-racial friendships. This makes it clear that there is a second driving force of ethnic friendship segregation over and above homophily in tie *creation*, namely homophily in tie *maintenance*. Thus, cross-sectional ethnic friendship segregation is in fact a snapshot of a dynamic process, in which both these mechanisms may play a major role. The main aim of this study is to analyse ethnic

homophily, differentiating between tie creation and tie maintenance.

Investigating those two driving forces of ethnic friendship segregation, a measure of which ties exist within and between groups, we study group-wise differences in the creation and dissolution of between- and within-group friendships over time. Because reciprocity – the propensity that someone that I regard as a friend also regards me as a friend – has been found to be a strong determinant of friendship maintenance (e.g. Rivera et al. 2010) we further analyse to what extent group differences in reciprocity play a role for homophily of friendship creation and maintenance.

In order to differentiate network segregation as a cross-sectional measure from homophily as a process, such that the first is an intermediary state of the second, we need longitudinal network data, which is a rare species. We draw on newly collected sociometric data from the recent cross-national school-based study Children of Immigrants Longitudinal Survey in four European countries (CILS4EU), which provides information on almost 19,000 pupils observed at age 14-15 (Kalter et al. 2016; 2018). We use data from Sweden, where network information is most attuned to our analytical set-up, comprising friendship networks from 194 classrooms, and studied for two consecutive years.

We fit random-coefficient multilevel network models to estimate pupils' tendency for ethnic homophily of tie creation and maintenance. We apply hierarchical stochastic actor oriented models (HSAOM) to test (i) whether pupils rather make friends with in-group classmates; (ii) rather stay friends with them than with those of different migration background, and (iii) whether this tendency is moderated by these friendships being reciprocated.

3.2 Theory

3.2.1 Theoretical outline: segregation, homophily and network mechanisms

This work understands ethnic segregation of adolescents friendship networks in classrooms as a macro-level phenomenon, the emergence and reproduction of which depend on micro-level actions and choices, much in line with the tenet in analytical sociology (Coleman 1990; Hedström 2005). In our approach, one of the micro-level mechanisms leading to segregated networks is ethnic homophily (McPherson et al 2001); individuals‘ tendency to befriend similar others. Ethnic homophily in adolescents friendship networks has received increasing attention (e.g. Munniksma et al. 2013; Smith et al. 2014; Smith et al. 2016; Stark 2015, Leszczensky and Pink 2015, Boda and Neray 2015, Jugert et al 2018; Leszczensky and Pink 2019, Wittek et al 2019), especially with regard to the role of ethnic or host-country identification or attitudes towards respective out-groups. Although analysing longitudinal network data has become more common (e.g. Leszczensky and Pink 2015, Boda and Neray 2015, Jugert et al 2018; Leszczensky and Pink 2019, Wittek et al 2019) homophily of friendship creation and maintenance has so far rarely been differentiated (Aboud et al. 2003; Rude and Herda 2010). While homophily in friendship formation can result in the emergence and increase of segregation, homophily of friendship maintenance can explain stability of segregation, even when friendships are predominantly created within groups.

Moreover, as has been demonstrated by Grund (2014), underlying network processes such as transitivity – the tendency to befriend the friends of one’s friends – can actually lead to segregated networks by amplifying only small tendencies of homophily in the early stages of a networks' evolution. One network process that might play an important role for homophily in friendship creation is reciprocity (Rude and Herda 2010; Vaquera and Kao 2008). We argue that to better understand the processes leading to segregated

(friendship) networks, it is important to analyse how networks evolve over time, and in doing so, to investigate group-specific tendencies of both tie creation and maintenance. Especially with respect to tie maintenance only little research has been conducted investigating the role of network-inherent processes, but reciprocity may well be a relevant predictor of friendship maintenance (Gershman and Hayes 1983). Indeed, it can be an additional force driving the segregation of a network if it affects the creation and maintenance of within- and between-group friendships differently.

3.2.2 Homophily of tie creation and maintenance

Most research on adolescent social networks has focused on the context of classrooms in schools because they define clear network boundaries and pupils are exposed to each other for a large proportion of the day, and often for many years. This provides them with plenty of opportunities for interactions in which pupils get to know each other and can form friendships[22]. The process of friendship formation requires actors to invest in the friendship in terms of material costs (e.g. financial resources for shared activities) as well as emotional and cognitive inputs. Once formed, the outcome of a friendship provides the actors with returns such as social confirmation, company, shared experiences, and emotional support. Students commit to investments into a friendship when they expect it to outweigh the costs (*friendship creation*). They continue to invest into the relation when their expectations are met to a satisfying extent (*friendship maintenance*). This is more likely when both actors invest into a friendship by *reciprocating* it, whereas unbalanced or one-sided friendships are more costly for one of the involved actors, making it more likely for the investing actor to stop and dissolve their tie. The related expectations about whether an investment is worthwhile come with

22 Of course, these opportunities are restricted by the composition of the classrooms – the opportunity structure, which in turn can be related to higher or lower levels of segregation (e.g. Smith et al 2016). Nonetheless, a systematic investigation of the opportunity structure would go beyond the scope of the work presented here.

respective uncertainties based on the perception of the other (*interpersonal uncertainties*).

As stated in group identity theory, adolescents tend to identify themselves with groups, and ethnic groups are an important reference characteristic for such group association (Taijfel and Turner, 1979). We argue that homophily (to some extent) arises from pupils' associating higher costs or having higher uncertainty when considering friendship with an out-group member compared to an in-group member (*intergroup uncertainties*). These might be influenced by existing prejudice and attitudes pupils hold about the respective out-group, thus showing an impact on pupils group-specific friendship choices (e.g. Schachner et al. 2015). Costs could also be sanctions from third parties, such as parents or other friends, who may be prejudiced against out-groups. When initiating friendships, students consequently expect lower investment costs for ethnic in-group classmates and greater uncertainty about receiving benefits from pupils of a different ethnic group. Therefore, we expect to find a tendency of pupils' friendship ties to be created more often to in-group members than out-group members (**H1a**).

Regarding the stability of friendships, Hallinan and Williams (1987) point to the attractiveness of alternatives and the friendship quality. Based on the points made on the quality of interactions in social relationships, one could argue that inter-ethnic relations firstly have to overcome visible differences of individuals' characteristics such as salient ethnicity. After this hurdle has been overcome, relationships may be more strongly founded on what could be called secondary homophily of friendship-strengthening characteristics (e.g., common interests or shared attitudes). However, Rude and Herda (2010) find that intragroup friendships were maintained longer. This might be the case if less-visible characteristics that pupils learn about each other, such as attitudes, opinions or interests, show stronger differences between groups. It might also be the case that

inter-group friendships are put under stronger strain from other in-group peers or parents and therefore are dissolved to reduce cognitive dissonance or emotional stress. Combining both arguments, salient differences between pupils of different background might have been overcome but less salient differences between groups might show later on. Therefore we expect to find that friendships within groups are more likely to be maintained than friendships between groups (*maintenance homophily*) but that this difference is lower than for friendship creation (**H1b**).

3.2.3 Reciprocity and homophily

Reciprocity is a common and important factor for the creation and maintenance of friendship ties. Reciprocated friendships are rewarding because they tend to involve mutual and reinforcing benefits, including trust and intimacy. Non-reciprocated ties (which may not even be seen as friendships for others than for ego) may on the other hand imply cognitive dissonance, emotional strain, or frustration for the person creating or maintaining a one-sided tie (and perhaps bad conciousness for the other). A widely-used explanation for the importance of reciprocity relies on social exchange theory (Emerson 1979, Rusbult and Buunk 1993). When people invest in friendships, they expect to receive similar returns (e.g. emotional support). If investments and returns are unbalanced, it is likely that an alternative is sought. Still, there might be situations which are not perceived as unbalanced but where non-reciprocated friendships persist. For example, this might be the case for dyads between pupils of different social status, where the student not reciprocating a friendship nomination has a higher status (e.g. Davis 1972). Pupils of low status might maintain such a friendship tie because the lack of returns from their investment in terms of reciprocity is compensated by feeling associated with a pupil of higher social status. Relating this line of reasoning to inter-ethnic friendships, it can be argued that native students on average hold higher social

status than non-natives. From this it would follow that non-reciprocated ties are more likely to be created and maintained between groups, notably from minority groups to the majority.

While a case could thus be made for assuming that non-reciprocated ties would be more common between than within groups, there is also reason to expect the opposite. Block (2015), argues that this is because within-group ties often present a low-cost situation. In his case, this is represented by a triadic network structure as compared to a high-cost (non-embedded) dyad. We can apply a similar logic to our ethnic background groups, following our arguments above: within-group dyads only come with inter-*personal* uncertainties whereas between-group dyads are additionally affected by inter-*group* uncertainties, implying that the former case is a low-cost situation compared to the latter. For example, while pupils from the same ethnic group might not be sure if they enjoy spending leisure time with another, pupils from different groups might additionally be unsure about shared cultural values. We therefore assume that the degree of homophily is greater for non-reciprocated rather than reciprocated ties. But how does such regularity depend on tie creation and dissolving, respectively?

First, because of the low-cost situation, we believe that one-directional ties are more likely to be created within rather than between groups, so we expect to find (strong) ethnic homophily in one-sided tie creations (**H2a**). On the other hand, reciprocity is more important for the creation of ties between than within groups so we expect to find lower levels of ethnic homophily for reciprocated ties (**H2b**). Because reciprocated friendships are more beneficial for both actors over the course of time, these friendships should also be expected to be maintained longer in general, whereas one-directed ties should be more easily dissolved. As within-group dyads pose lower costs for the involved actors they have a higher chance of being preserved than between-group

dyads, so we expect to find homophily in the maintenance of one-sided friendships (**H2c**). Reciprocity should therefore be more relevant between groups than within groups, so again we expect to find lower levels of maintenance homophily (or even heterophily) for reciprocal friendships (**H2d**).

3.3 Data and Variables

We analyse a sample of 194 Swedish classrooms first observed at age 14-15 and the subsequent wave about one year later (Kalter et al. 2016; 2018). School classrooms pose an ideal context to study friendship networks since pupils spend time with each other for large proportions of the day, thus making informed choices about whom to become or not to become friends with. The first wave of the Swedish data was collected in grade 8, when some classes had been formed the year before, and others normally 2-4 years before, and the follow-up was conducted a year after that. Our dependent network is constructed from pupils nominating their five best friends in their class. This results in a directed network with a maximum out-degree of 5.

We differentiate pupils as being either native or being first or second generation migrants. The migration to Sweden is large (around 15% are foreign-born) and heterogeneous. The majority of our respondents come from a non-European background, particularly the Middle East; other large groups originate in ex-Yugoslavia, Africa and Eastern Europe. We consider those who either immigrated themselves or who have two immigrated parents (or one, in case of lone parents) as having an immigrant background. A more detailed classification is not possible because of the variety of immigrant groups, leading to small n at the classroom level.

Since we want to investigate ethnic homophily, net of homophily on another level, we are conditioning on a number of personal traits that have been found to be significant

predictors of friendship (McPherson et al. 2001). Based on this, we aim at controlling for friendship selection on the basis of demographic and behavioural characteristics.

Gender homophily is a tendency that is found in virtually all adolescent friendship networks (e.g. Shrum et al. 1988); we therefore expect this to play an important role in our friendship networks as well. It has furthermore been found that students with similar cognitive and language ability tend to become friends (e.g. Almark 1922; Richardson 1940). Besides the two direct controls for this, we include age as well. Although students in the same class are of relatively similar age anyway, there are differences to be expected when it comes to grade repeaters, which can be seen as another measure of cognitive ability. At the same time, it is tapping into the dimension of how seriously school work is taken, which is a likely predictor of friendship as well, relating to shared aspirations and goal orientation (Duncan et al. 1968). Being of a similar socio-economic status is another likely predictor for friendship (e.g. Bott 1929; Loomis 1946). It is moreover likely that friends share aspirations and cultural tastes (Kalmijn 1998, Marsden & Gorman 2001), as well as similar financial resources (Hjalmarsson & Mood 2015). This is accounted for in two ways: through similarity in parental occupational prestige, as well as traditional family structure, i.e. whether both biological parents live together with the pupil. Furthermore, similar smoking behaviour has been repeatedly found to be a significant predictor of friendship selection (e.g. Cohen 1977; Kandel 1978).

We furthermore control for the tendency of a pupil to nominate fewer friends in the class when he or she has more friends outside of the school class, to capture baseline differences in nominating behaviour.

3.3.1 Variable definitions

Gender: Based on self-reports in the questionnaire (boy/girl).

Cognitive ability: The cognitive test used is based on visual puzzles (‘patterns’), minimising cultural bias (Weiß, 2006). It is aimed at measuring fluid intelligence, the main component of general intelligence. The test comprised 27 items and was clocked at 7 minutes; the resulting test scores were z-standardised.

Language ability: The language ability test is based on a test of antonyms, and z-standardised.

Occupational prestige (ISEI scale): ISEI is based on the coding of parents’ occupation into the 2008 International Standard Classification of Occupations (ISCO-08) and converted into the interval-scale ISEI-08 occupational status (Ganzeboom and Treiman, 1996).

Smoking: This binary variable accounts for whether the individual smokes. This is based on a categorical variable with five values: “every day”, “once or several times a week”, “once or several times a month”, “less often”, and “never”. The respondent is considered to be a smoker, if he or she named a category other then “never”.

Traditional family structure: This is a dummy variable accounting for whether both biological parents live with the pupil.

Outside friends: This variable accounts for the number of friends outside the school class. It is based on ego-centric network questions on the five best friends in general, in which the respondent answered questions about his or her five best friends without identifying them directly. One question asks whether the friend is in the same school class as the respondent.

3.3.2 Descriptives

Table 3.1 shows the descriptives for the 194 networks[23] in our sample. The classrooms size ranges between 13 and 34 pupils, with an average of about 23, and have between 3% and 87% non-native pupils with an average of 31%. The average class consists of about 51% percent females, 14% pupils who smoke and 65% coming from a traditional family structure. Pupils were almost 15 years old at the first wave of data collection.

The average outdegree within classes slightly reduces between waves from 2.92 to 2.73, which is reflected also in the densities – that is, the proportion of all possible ties that are realized – which reduce from .18 to .17. This implies that slightly more ties are dissolved than created between waves, showing the potential importance of tie maintenance for changes in classroom segregation.

23 From the original sample of 251 classes, 57 were excluded because they either consist of only native or only non-native students, have an average outdegree lower than 1.5 or higher than 15, or too little or too high change between waves (indicated by a jaccard coefficient – measuring the proportion of ties that are stable over the number of ties that exist in either or both data collection waves – below 0.3 or above 0.98). Since the SAOMs model changes that actors make, networks with too little or too much change could indicate too small cell counts resulting in problems for the estimation procedure.

Actor variables – Classroom aggregates	min	mean	max	sd
N pupils	13	22.71	34	3.86
% non-native	.03	.31	.96	.26
% female	.17	.51	.87	.13
% smoking	0	.14	.62	.11
% traditional family	.35	.65	.94	.12
age	14.32	14.64	15.22	.15
ISEI	29.31	46.08	65.60	8.71
language ability	11.26	18.41	23.16	2.13
cognitive ability	12.60	17.77	21.95	1.87

Friendship Network	Wave 1				Wave 2			
	min	mean	max	sd	min	mean	max	sd
average outdegree	1.39	2.92	4.286	.55	1.62	2.73	4.44	.52
density	.11	.18	.34	.04	.10	.17	.43	.05

N classes = 189

Table 3.1: Basic descriptives

3.4 Analytical approach and Method

In order to gain some initial insights about our hypotheses, we first look at descriptive statistics of group-wise creation and maintenance for unbalanced and reciprocal ties. Following that, we estimate models in order to assess ethnic friendship selection while controlling for other homophily effects. Estimates are obtained from applying random-coefficient multilevel network models (Koskinen and Snijders, 2018). In the classical Stochastic Actor-Oriented Model (SAOM) the evolution of networks is modelled in continuous time, in which actors change their networks. The changes actors make are so called *mini-steps* and are the basis for inference. The *mini-step* changes refer to individuals' choices of either creating a tie, maintaining a tie, or doing nothing. These changes are modelled in the *evaluation* function, which indicates tendencies of stability or change of respective network conditions (e.g. transitive triplets). More precisely, estimated parameters of the evaluation function can be interpreted as multinomial log-odds ratios of actors creating or dissolving ties versus not changing their current network.

The model also allows to further specify parameters for tie creation, which estimate effects for creating ties compared to not creating ties. If these parameters are included additionally to evaluation effects, the latter represent tie maintenance and the linear combination of both refers to effects of tie creation. As for our case, having longitudinal data on several network groups, i.e, multiple classrooms, the HSAOM model allows for heterogeneity between the different classrooms (Koskinen and Snijders forthcoming). The HSAOM can be understood with reference to classical multilevel regression type models as they differentiate individual- (pupils) and group-level (classroom-networks) variation.

While all classrooms are fitted with the same SAOM, effects specified as varying have a specific set of parameters $\theta^{[g]}$ for each classroom (g). A multivariate normal distribution with mean μ and covariance σ forms the assumed population-level model from which parameters are drawn, with the inference target being the population mean (or global mean in Bayesian terms) rather than the classroom specific parameters. During the estimation, samples of three core posterior variates are drawn through iterative Markov-chain Monte Carlo (MCMC) sampling, which relate to the population parameters, the classroom-level parameters, and the unobserved changes between waves. Since inference is based on the posterior draws given observed data, it is important to make sure that the MCMC samples are efficient. Our analyses are based on a first run with 300 main phase iterations which are used as starting values for a following 8,000 main phase iterations (Ripley et al., 2018)[24].

In our analytical strategy we first estimate a model including a parameter accounting for ethnic homophily for the creation (**H1a**) and evaluation (**H1b**) function. Setting up the model in this way, a positive homophily parameter on the evaluation function shows a tendency for friendships to rather be maintained within groups than between. Friendship creation, on the other hand, is represented by the linear combination of the evaluation and creation parameters for homophily. Thus, if the parameters for homophily in the evaluation and creation function are positive, we find support for H1a and H1b, with the evaluation homophily indicating homophily of friendship maintenance and the creation homophily indicating stronger homophily of friendship creation.

The second model adds group-specific reciprocity for tie creation (**H2b**) and maintenance (**H2d**), such that the initial ethnic homophily effect now only refers to non-reciprocated tie creation (**H2a**) and maintenance (**H2c**). Here, the homophily parameter

24 Diagnostics of the MCMC convergence indicate that the length of the main-phase algorithm is sufficient (not shown, but available on request).

of the evaluation function represents one-sided friendship maintenance, and supports our hypothesis (**H2c**) if found to be positive. The interaction between homophily and reciprocity in the evaluation function supports our hypothesis (**H2d**) if it shows to be negative, indicating homophily of reciprocated friendship maintenance to be lower than homophily of one-sided friendship maintenance. *Heterophily* of reciprocated friendship maintenance would be indicated if the linear combination of the evaluation parameters for homophily and its interaction with reciprocity are significantly less than zero. The homophily effect of the creation function supports homophily of one-sided friendship creation (**H2a**) if its linear combination with the homophily effect of the evaluation function is positive, which is the case if both parameters by themselves are positively significant. Our expectation about finding lower levels of reciprocal friendship creation (**H2b**) is met if the interaction effect for homophily and reciprocity in the creation function is negative.

All models include the same set of endogenous network effects in order to improve the simulation of network changes as well as to ensure that the specified effects of interest actually represent the network situations hypothesised about. The included endogenous effects account for transitive triples (the tendency for ego to create and maintain ties to friends of friends), reciprocated transitive triplets (the tendency for ego to reciprocate ties from friends of friends), indegree popularity (the tendency of egos to create and maintain ties to alters that are popular in the network), outdegree popularity (the tendency of egos to create and maintain ties to alters that have ties to many others), outdegree activity (the tendency of egos to nominate even more alters when already having many ties to others) and the number of existing friends outside of the classroom.

Moreover, we control for several potential alternative dimensions of homophily besides migration background, including parameters to model homophily of gender, smoking,

parental social background, age, language ability, cognitive ability as well as similarity in family background. The endogenoeus network effects as well as other homophily dimensions are specified in the evaluation function, thus being taken into account in the estimation of ethnic homophily of creation and maintenance, although not explicitly modelled. All parameters are specified as random, with the exception of similarity in cognitive ability, because its variance parameter showed problems with convergence.

3.5 Results

Figure 3.1 shows segregation measured on three scales, calculated for each network for students that participated in both waves. The EI-index is calculated as the existing ties between groups minus the in-group ties as a proportion of all existing ties. Here the sign was reversed so that positive numbers indicate segregation and shows high levels of segregation for most of the classrooms. In contrast, Yules' Q indicates that most classrooms have low levels of segregation. It could be argued that the absence of segregation for a majority of the classrooms makes the investigation of homophily unnecessary. However, because homophily is a mechanism that works over time, desegregated classrooms may well change towards a more segregated state due to pupils' homophileous friendship choices. Why do our applied measures provide contradictory results? The difference is that Q is invariant for differences in group size – by controlling for the margins, Q shows the propensity for ties given the opportunity structure, in this case, the presence of alters with a different ethnicity. Thus, the high segregation indicated by the EI-index shows that more ties exist within than between groups and no segregation revealed by Q shows that the first is due to differences in group size.

For a more intuitive measure, we introduce differences in group-wise densities, calculated as the difference between the proportion of realised in-group ties and possible

in-group ties minus the corresponding difference between the proportion of realised between-group ties and possible between-group ties. This measure accounts for relative group-sizes and is easily interpretable. It becomes one if all possible in-group ties exist but no ties between-groups, and minus one for the opposite case.

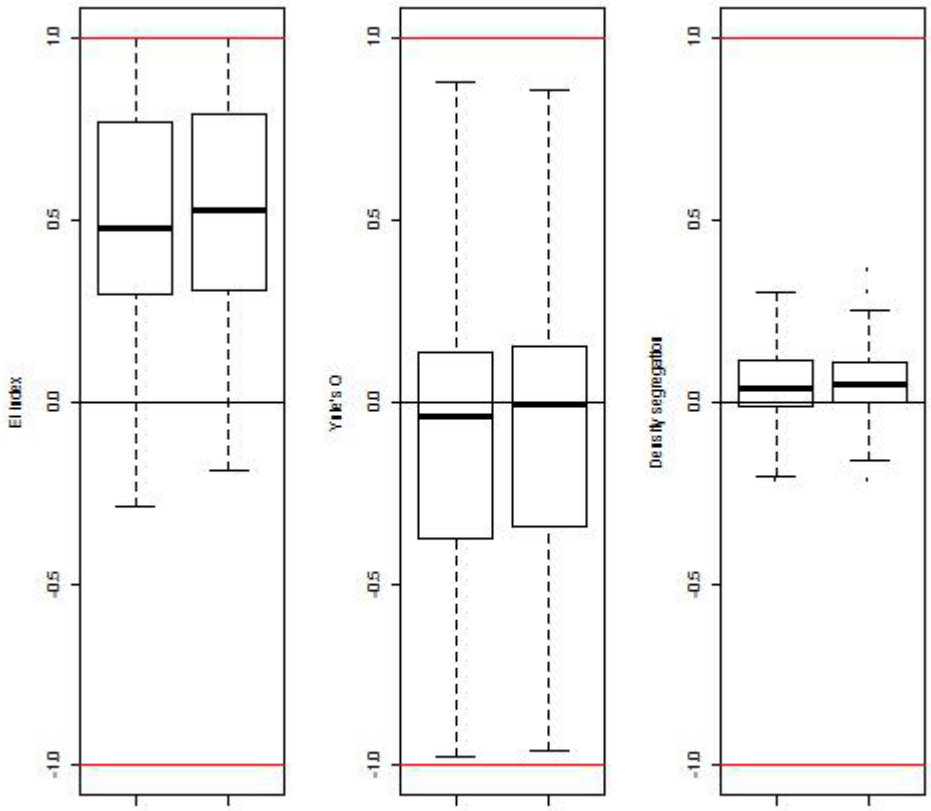


Figure 3.1: Measures of segregation at both waves

The density differences indicate for most classes a moderate level of segregation, between 5 to 10 percent more realised in-group than between-group ties. However, all three measures have in common that they show only little difference in the distribution of segregation between waves despite changing ties in the networks. In order to illustrate the dynamics of segregation, figure 3.2 shows the change of each networks' level of segregation. The y-axis shows the segregation based on groupwise densities as

mentioned above. The x-axis plots the rank of segregation in each classroom during the first wave, with the classes that have the lowest-segregation to the left and the highest to the right. The black line in the centre shows the segregation value at the first observation, red bars indicate an increase of segregation and blue bars a reduction. It is clear that segregation levels of individual networks actually change considerably between observations, although the distribution of segregation between networks remains stable. Figure 3.2 also reveals a tendency towards the mean, meaning that classes starting with lower segregation more often increase, while those with initially higher segregation tend to decrease.

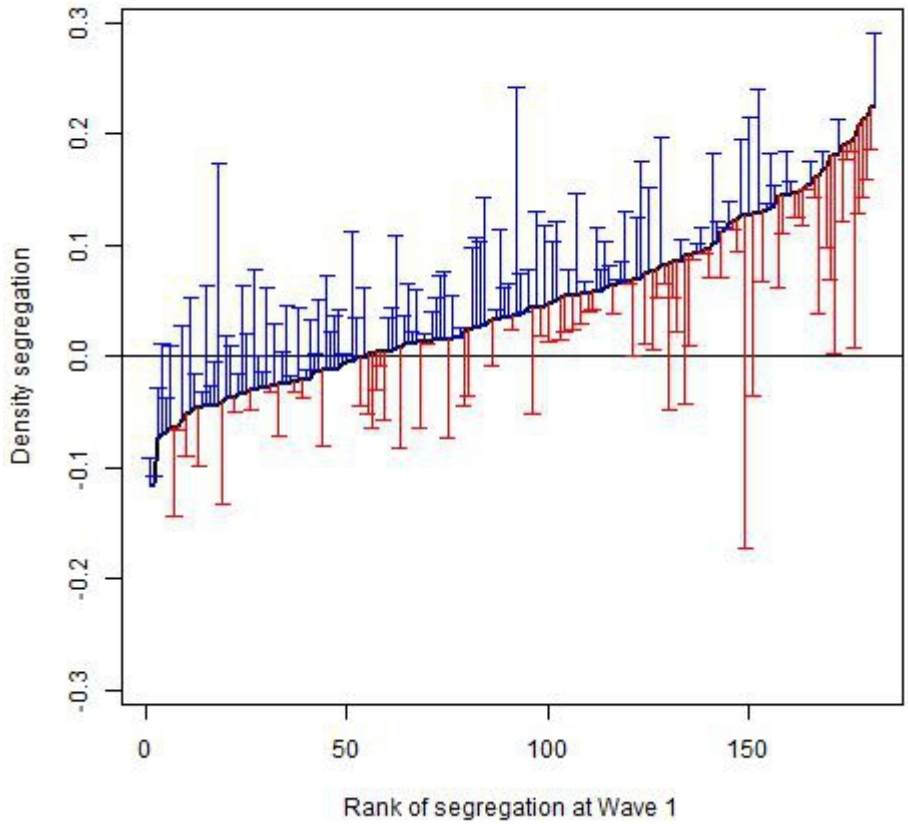


Figure 3.2: Change of segregation in classrooms between waves

In order to better understand how individual actors decide about changes in their

network, and to get a first indication for how our hypotheses fare, we next descriptively investigate creation and maintenance of ties, differentiated for uni-directional and reciprocated ties.

Figure 3.3 shows statistics of tie changes within and between groups. The upper box provides information on tie creation and the lower on tie maintenance. The left hand side in the box provides an illustration of the respective tie change under investigation, showing the relation between two pupils at wave 1 and wave 2. The upper part in a box (above the dashed line) looks at one-sided friendships, and the lower at reciprocated relationships. For reciprocated friendship creation, the dashed lines for wave 1 indicate that reciprocation can occur either if a pupil has already been nominated at wave 1 and thus creates one tie to reciprocate the initial nomination, or no ties exist at wave 1 and two reciprocal ties are created. In the lower part for friendship maintenance, the dashed lines between pupils at wave 1 indicate the ties that are present at wave 1 and continue to be present at wave 2. The rows with one white and one grey circle refer to relationships between pupils of different migration background, and the rows with two white circles refer to both pupils having the same background.

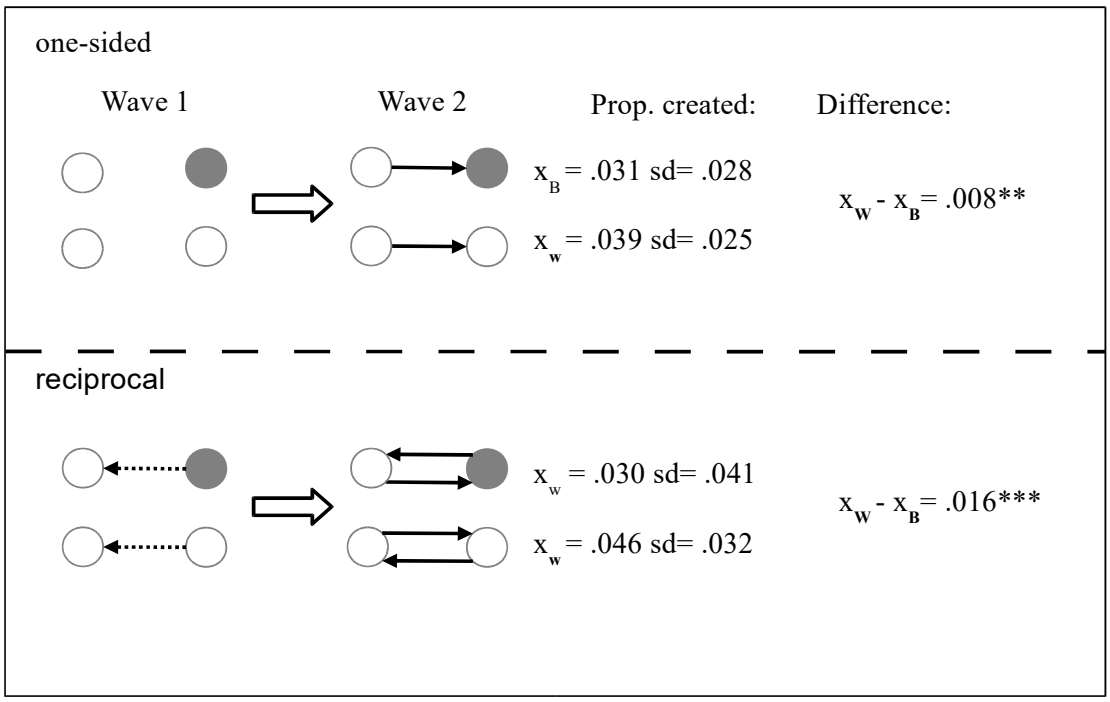
The third column shows the proportion of ties that are either created from all possible ties that can be created, or the ties that are maintained from those that are already present at wave 1. These proportions are reported as average and standard deviation over the classrooms in the sample for the respective within- and between-group dyads. The fourth column reports the average difference of subtracting the proportion of ties within minus ties between groups, stars indicating whether the between and within means differ in a t-test.

The results, in the upper part in figure 3.3 (rightmost part), show that on average less than one percent more uni-directional in-group ties are realised within than between

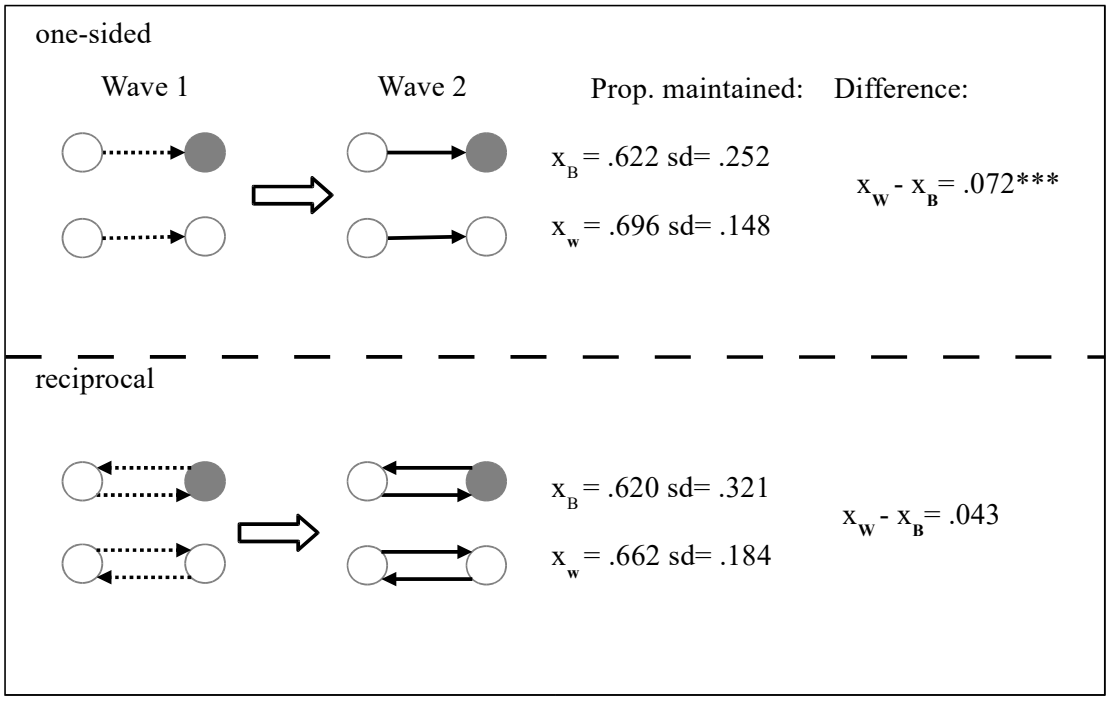
groups, with the difference being statistically significant ($t = 3.119$, $p < .01$) supporting **H2a**. With regard to reciprocated ties, this difference is even stronger with 1.6 percent more in-group ties being realised than between-group ties ($t=3.917$, $p<0.001$), which is not in line with the expectations from **H2b**.

The bottom box of figure 3.3 shows the calculated statistics for tie maintenance. About seven percent more uni-directional in-group ties are preserved than between-group ties ($t=3.331$, $p < 0.001$), in line with **H2c**. In contrast, the maintenance of reciprocated ties shows the smallest difference, with only 4.3 percent more in-group ties being preserved than out-group ties ($t=1.517$, $p=.135$) supporting **H2d**. In sum, investigating the tie changes between waves shows one-directional and reciprocal homophily for tie creation and homophily for one-sided tie maintenance but no homophily for reciprocated tie maintenance.

Homophily of friendship c**reation:** One-sided and reciprocated



Homophily of friendship maintenance**:** One-sided and reciprocated



T-Test of mean differences: ** p<0.01 *** p<0.001

Figure 3.3. Homophily of friendship creation and maintenance by reciprocity

The results of our hypotheses testing of homophily, while controlling for friendship

selectivity, can be found in Table 3.2.[25] The first column (**M1**) indicates support for ethnic homophily with regard to tie creation (**H1a**) and maintenance (**H1b**) with the average effects being positively significant (b_{1a} = .587, se= .104; b_{1b}=.241, se=.133). The respective odds ratios show that ties have a 1.8 times higher odds to be created rather than not, and 1.3 higher odds to be maintained than dissolved within groups than between. Thus, pupils tend to generally create and maintain ties to other pupils that have the same migration background.

The second column shows the model differentiating tie creation between uni-directional and reciprocated ties (**M2**). It shows support for one-sided homophily of tie creation (**H2a**) with the effect being positively significant (b2a=.619, se=.226), indicating that non-reciprocated ties have a 1.85 higher odds to be created within groups than between. Pupils create non-reciprocated friendship ties to others of their own group rather than to those who have a different migration background.

On the other hand, the model provides no support for reciprocated homophily of tie creation to be lower than assymetric homophily (**H2b**) with the respective interaction not being significant (b2b=.108, se=.274). Pupils have a tendency to rather befriend those with the same than those with a different migration background, irrespective of these friendship being mutual or not.

With regard to tie maintenance, we find a positive (and significant) effect for ethnic homophily of uni-directional friendships (b2c=.599, se=.217), supporting the expectation that assymetric friendships have a higher chance to persist within groups than between (**H2c)**. Finally, this model indicates strong support for lower homophily of friendship maintenance of reciprocated (rather than non-reciprocated) ties (**H2d)** with the respective interaction effect being significantly negative (b2d= -.844, se=.230). So,

25 The table presents linear combinations that represent the comparisons of effects referring to the hypotheses. The full table of estimated parametres can be found in Appendix Table 1.

one-sided friendships have a 1.8 times higher odds to be preserved within groups than between. However, reciprocated friendships have 0.8 times lower chances to be maintained within compared to between groups, although this difference fails to reach conventional significance levels (b= -.247, se=.184).

Effect names		M1		M2	
		b	s.e.	b	s.e.
Maintenance					
Core effects					
	Reciprocity	3.339	.433 ***	4.265	.572 ***
	Homophily	.241	.133 *	.599	.217 ***
	Reciprocity **X** Homophily			-.844	.230 ***
Creation †					
Core effects (creation)					
	Reciprocity	4.385	.560 ***	4.988	.978 ***
	Homophily	.587	.104 ***	.619	.226 **
	Reciprocity **X** Homophily			.108	.274
Variance parameters					
Maintenance					
	Reciprocity	2.013	.530	2.707	1.268
	Homophily	1.269	.291	1.417	.377
	Reciprocity **X** Homophily			1.719	.489
Creation					
	Reciprocity	2.702	.748	3.537	1.331
	Homophily	1.604	.410	1.742	.481
	Reciprocity **X** Homophily			2.263	.854

***Distribution of parameters does not cover 0 ** Less then .01 of distribution covers 0

† Paramters are a linear combination of respective evaluation and creation effects

Table 3.2: Results of Hierarchical SAOM

3.6 Conclusion and discussion

Our analysis finds evidence for general ethnic homophily of tie creation and

maintenance, thereby supporting the first (**H1a**) and second (**H1b**) hypothesis. When differentiating one-sided from reciprocal friendships, we find homophily of one-sided friendship creation (**H2a)** but not that the reciprocal case is weaker (**H2b**). However, with regard to tie maintenance we find homophily of one-sided friendship maintenance (**H2c**) and heterophily of reciprocated friendship maintenance (**H2d**).

When understanding segregation as a macro-phenomenon that can be observed in snapshots of a social network, one of its driving forces is the individual-level mechanism homophily. Individuals have a tendency to connect to similar others, in our case pupils rather befriend those of the same migration background. In an emerging network, for example a school classroom where pupil are recently put together, actors' tendency to make friends with similar others can lead to strong segregation (and not only on ethnic grounds). However, to understand the dynamics of segregation over time, it should be considered that homophily can also express itself in the form of tie maintenance; pupils having a tendency to stay friends with similar others. Moreover, both of these aspects can be interrelated with endogeneous network mechanisms, such as reciprocity.

Our finding that ethnic homophily initially seems weaker for tie maintenance, changes when differentiating between one-sided and reciprocal relationships. This implies that friendships within and between groups, when not reciprocated, might refer to different social understandings of the involved actors, while reciprocated friendships seemingly have similar dynamics. It becomes clear that homophily of tie maintenance is strong for one-sided friendships but does not exist for reciprocal friendships. Thus, once a balanced friendship relation is established, pupils keep these whether their friends are of the same or a different immigrant background. This supports the argument that inter-ethnic differences are overcome once a relationship to someone in an out-group is

established.

Finding homophily with regard to the maintenance of unbalanced friendships raises the question about the nature of such relations: Why is it that inter-ethnic, as compared with in-group nominations cease more often when not reciprocated? One reason could be that less salient characteristics like cultural or socioeconomic differences are greater between ethnic groups than within, and such differences are likely to create strain in relationships, at least at some point. One-sided friendships being more fragile than reciprocated ones means that they might rather be dissolved under the additional strain of group differences, but also under the influence of third parties, for example if other friends or relatives hold prejudices against inter-ethnic friendships.

While we find homophily of friendship creation, this does not differ between one-sided and mutual friendships. One possible explanation is that a pupil choosing to nominate another as their friend does not depend on them being nominated in return because they are not aware of the alters nomination at that point. However, over the course of such a friendship qualitative differences might become apparent to pupils such that maintenance of these might differ whether they are mutual or one-sided. From our results it appears that friendships are more likely to be created within groups, though equally so for one-sided and mutual relations. This means that more ties are established within groups which contributes to an increase of segregation. This tendency would be amplified if one-sided ties are rather dissolved between groups providing even stronger tendencies for increasing segregation.

We also have to entertain the possibility that we underestimate homophily in tie creation. Because our respondents have spent at least one year together, and often more than so, social networks are already established and visible characteristics such as different migration backgrounds might play a weaker role since pupils might already

know more about each others non-visible characteristics. If we would have had network data on randomly assembled adolescents, tie creation (and homophily in such ties) would probably have been more common, due to salient differences playing a stronger role of differentiation.

On the other hand, relating Feld's theory on social *foci* (1981) one can argue that it is unlikely to find randomly assembled classrooms where pupils have no previous relations with each other. Due to housing segregation and access to schools being dependent on living area, pupils that end up in the same classroom have a higher probability to come from the same neighbourhood. It is therefore likely that pupils already are part of other *foci* such as sport clubs or religion-based communities where they have met and gotten to know each other. The more *foci* they share, the more positively they mutually associate the relationship with each other, thus being more likely to become friends. In line with this, homophily is rather the result of housing segregation and shared association with subsequent participation in social *foci* which are segregated as well. In this case, analysing homophily in school classrooms reflects pupils of similar background having more shared *foci* due to housing segregation.

Put in this light, our finding of reciprocated friendships being more likely to be maintained between pupils with different backgrounds (compared to those with a shared background) could be due to at least two explanations. Firstly, in our sample these mixed dyads share specific *foci* that provide stronger positive associations between both pupils and which more rarely occur for pupils of the same migration background. Or secondly, in those *foci* which are shared by pupils of different backgrounds, their relations are more positively ascribed than those to pupils with the same background. Either way, our research contributes by revealing this phenomenon, which can help to counter segregation. For this to be the case, future research should a) focus on a better

understanding of what sets these reciprocated friendships of pupils from different backgrounds apart from those between pupils of shared background and b) identify the ways in which these friendships are established in order to provide better opportunities for them to emerge.

For high immigration societies, our work gives some insight into mechanisms counteracting social integration, namely homophily of friendship creation and maintenance. We find that pupils tend to rather become friends with others that have the same migration background, which can create socially segregated groups of natives or non-natives who mainly dwell in their own circles. It can be argued that looking at early stages of secondary school does not necessarily represent the "natural" emergence of such friendship networks, especially considering housing segregation and subsequent opportunity structures for interactions (e.g. Feld 1981).

Additionally considering friendship maintenance, the dynamics of pupils friendship networks are most likely not constant and might not even be linear, depending on the age of a network. Networks where pupils just got to meet each other are more likely to be driven by friendship creation, because fewer friendships exist that can be maintained; the network first has to emerge. Only later on, the maintenance of friendships can play a stronger role. Thus, we can not claim that the mechanisms observed in our sample of Swedish classrooms are the cause of segregated classrooms, because it might be a result of pupils previous interactions in other *foci*. However, investigating how pupils changed their friendship choices after a year, rather than looking at a single snapshot of these networks, allowed us to identify important differences in friendship creation and maintenance. The fact that our data did not show network growth (i.e. pupils having in average more friends at the later observation), but rather a slight tendency for the opposite, supports the importance of differentiating friendship creation and

maintenance. While pupils on average nominate about one new friend, they have already nearly three friendships to consider maintaining. Nearly 1 more percent of the created friendships are towards pupils of the same group, but about 8 more percent of the friendships with the own groups are maintained than friendships with pupils from the other group. These numbers might appear small, especially considering only a single friendship network in a classroom, but over the course of several school years they may set off a non-negligible trend towards segregation. Not only do pupils tend to choose their new friends from their own group, they more often choose to stay friends within this group too. For the group of non-natives this means that they not only get slightly fewer native friends who could provide them with beneficial social capital, but that they have less access to those especially beneficial long lasting contacts. Our findings have relevant implications for subsequent research on the dynamics of friendship segregation, because we might underestimate tendencies of homophily when not differentiating reciprocated and one-sided in- and out-group ties. Moreover, it opens at least two relevant lines of research with regard to, firstly, why non-reciprocated between-group friendships are less stable and, secondly, potential group differences between native and non-native groups. In conclusion, our study showed that the investigation of (ethnic) homophily reveals relevant differences between one-sided and reciprocal friendships with regard to an in-group and between-group context. Additional findings indicate that observed segregation results to a strong degree on homophily of (one-sided) friendship maintenance and less on friendship creation.

3.7 Literature

Almack JC. (1922). The influence of intelligence on the selection of associates. *School and Society* 16:529-30.

Aboud, F., Mendelson, M., and. K. Purdy. (2003). Cross-Race Peer Relations and Friendship Quality. *International Journal of Behavioral Development.* 27(2), 165-

73.

Block, Per (2015): Reciprocity, transitivity, and the mysterious three-cycle. *Social Networks* 40, 163–173.

Blumer, H. (1958), Race Prejudice as a Sense of Group Position. *Pacific Sociological Review.* 1 (1), 37.

Boda, Z. and B. Néray. (2015). Inter-ethnic friendship and negative ties in secondary school. *Social Networks.*43. 57-72.

Bott H. (1928). Observation of play activities in a nursery school. *Genetic Psychology Monogr.aphs.* 4:44-88

Cohen J. (1977). Sources of peer group homogeneity. *Sociology of Education.* 50. 227-41.

Currarini, S., Jackson, M. O., and P. Pin. (2010). Identifying the Roles of Race-Based Choice and Chance in High School Friendship Network Formation. *Proceedings of the National Academy of Sciences.* 107(11). 4857-861.

Coleman, J. (1990). Foundations of social theory. *Harvard University Press*

Davis, J.A., Leinhardt,S. (1972). The structure of positive interpersonal relations in small groups. Joseph Berger (Ed.): *Sociological theories in progress*. Boston, Mass.: Houghton Mifflin (2).218–51.

Duncan O.D, Haller AO and A. Portes. (1968). Peer influences on aspirations: a reinterpretation. *American Journal of Sociology.* 74. 119-37

Emerson, R.M., 1976. Social exchange theory. *Annual Review of Sociology* 2, 335–362.

Feld, S. (1981). The Focused Organization of Social Ties. *American Journal of Sociology, 86*(5), 1015-35

Ganzeboom, H. B. and D.J. Treiman. (1996). Internationally Comparable Measures of Occupational Status for the 1988 International Standard Classification of Occupations, *Social Science Research*. (25). 201–39.

Gershman, E.S., & Hayes, D.S. (1983). Differential Stability of Reciprocal Friendships and Unilateral Relationships Among Preschool Children. *Merrill-Palmer Quarterly - Journal of Developmental Psychology*, 29, 169-77.

Granovetter, M. S. (1973). The Strength of Weak Ties. *American Journal of Sociology*. 78(6). 1360- 380.

Grund, T. (2014) Network Size and Homophily: Same-Sex Friendships in 595 Scandinavian Schools, in Manzo, G. (eds.) (2014) *Analytical Sociology. Norms, Actions, Networks*. New York: Wiley. 299-316.

Hallinan, M. T., and R.A. Williams. (1989). Interracial Friendship Choices in Secondary Schools. *American Sociological Review.* 54(1). 67-78.

Heath, A.F. and Y. Brinbaum (eds.) (2014). Unequal Attainments: Ethnic Educational Inequalities in Ten Western Countries. *Oxford: Oxford University Press*.

Heath, A. F., Rothon, C. and E. Kilpi. (2008). The Second Generation in Western Europe: Education, Unemployment, and Occupational Attainment. *Annual Review of Sociology. 34*(1), 211–35.

Hedström, P. (2005). Dissecting the social. On the principles of analytical sociology. Cambridge: *Cambridge University Press*.

Hjalmarsson, S., & Mood, C. (2015). Do poorer youth have fewer friends? The role of household and child economic reources in adolescent school-class friendships. *Children and Youth Services Review 57*, 201–11.

Jugert, P., Leszczensky, L., and Pink, S. (2018). The effects of ethnic minority adolescents' ethnic self-identification on friendship selection. *Journal of Research on Adolescence*, **28**, 379–95.

Kalmijn M. (1998). Intermarriage and homogamy: causes, patterns and trends. *Annual Review of Sociology.* 24:395-421.

Kalter, F., Heath, A.F., Hewstone, M. Jonsson, J.O., Kalmijn, M. Kogan. I., and F. van Tubergen. (2016) Children of Immigrants Longitudinal Survey in Four European Countries (CILS4EU) – Full version. Data file for on-site use. GESIS Data Archive, Cologne, ZA5353 Data file Version 1.2.0, doi:10.4232/cils4eu.5353.1.2.0.

Kalter, F., Jonsson, J.O., F. van Tubergen and Heath, A.F. (Eds.) (2018). *Growing up in Diverse Societies*. Oxford: Oxford University Press.

Kandel D.B. (1978). Homophily, selection and socialization in adolescent friendships. *American Journal of Sociology*. 84:427-36.

Koopmans, R. (2005). *Contested citizenship: immigration and cultural diversity in Europe. Minneapolis*: University of Minnesota Press.

Koskinen, J. and T.A.B. Snijders (2019). Random-Effecct Multilevel Stochastic Actor-Oriented Models (*in preparation*).

Leszczensky, L., and S. Pink. (2015). Ethnic Segregation of Friendship Networks in School: Testing a Rational-Choice Argument of Differences in Ethnic Homophily between Classroom- and Grade-Level Networks. *Social Networks.* 42. 18-26.

Leszczensky, L., and S. Pink (2019). What Drives Ethnic Homophily? A Relational Approach on How Ethnic Identification Moderates Preferences for Same-Ethnic Friends. *American Sociological Review*, 84(3). 394-419.

Lin, N. (1990). Social Resources and Social Mobility : A Structural Theory of Status Attainment. 247-71 in *Social Mobility and Social Structure*. Ed. Ronald L. Breiger. NY : Cambridge University Press.

Loomis CP. (1946). Political and occupational cleavages in a Hanoverian village. *Sociometry.* 9. 316-33.

Marsden P.V. and E.H. Gorman (2001). Social networks, job changes and recruitment. Sourcebook on Labor Markets: Evolving Structures and Processes, ed. I Berg, AL Kalleberg. New York: Kluwer Academic/ Plenum.

McPherson, M., Smith-Lovin, L., and M.J. Cook. (2001). Birds of a feather: Homophily in social networks. *Annual review of sociology*. 415-44.

Moody, J. (2001). Race, School Integration, and Friendship Segregation in America, *American Journal of Sociology* 107, no. 3: 679–716.

Munniksma, A., Stark, T.H., Verkuyten, M., Flache, A., and R. Veenstra (2013) Extended intergroup friendships within social settings: The moderating role of initial out-group attitudes. *Group Processes & Intergroup Relations.* 16(6). 752–70.

Pettigrew, T. F. and L. R. Tropp. (2006). A Meta-Analytic Test of Intergroup Contact Theory. *Journal of Personality.* 90(5). 751-83.

Richardson HM. (1940). Community of values as a factor in friendships of college and adult women. *Journal of Social Psychology.* 11:303-12.

Ripley, R.M., Snijders, T.A.B., Boda, Z., Vörös, A., and P. Preciado (2018). *Manual for SIENA version 4.0.* Oxford: University of Oxford, Department of Statistics; Nuffield College.

Rivera, M. T., S. B. Soderstrom, and B. Uzzi. (2010). Dynamics of Dyads in Social Networks: Assortative, Relational, and Proximity Mechanisms. *Annual Review of Sociology.* 36(1). 91–115.

Rude, J., and D. Herda. (2010). Best Friends Forever? Race and the Stability of Adolescent Friendships. *Social Forces*. 89(2). 585-607.

Rusbult, C.E., Buunk, B.P., (1993). Commitment processes in close relationships: an interdependence analysis? *Journal Social Personal Relations.* 10(2). 175–204.

Shrum W, Cheek NH Jr., Hunter SM. (1988). Friendship in school: gender and racial homophily. *Sociology of Education*. 61. 227-39.

Shore, C. (2000). *Building Europe: the cultural politics of European integration*. London: Routledge.

Schachner, M, K, Brenick, A., Noack P., Fons J.R. van de Vijver, and B. Heizmann. (2015). Structural and Normative Conditions for Interethnic Friendships in Multiethnic Classrooms. *International Journal of Intercultural Relations* 47. 1–12.

Smith, S. Maas, I. and F. v. Tubergen (2014) Ethnic ingroup friendships in schools: Testing the by-product hypothesis in England, Germany, the Netherlands and Sweden, *Social Networks*, 39, 33-45,

Smith, S., Van Tubergen F., Maas, I. and D.A. McFarland (2016). Ethnic composition and friendship segregation: Differential effects for adolescent natives and immigrants. *American Journal of Sociology.* 121 (4). 1223-72.

Stark, T. H. (2015). Understanding the Selection Bias - Social Network Processes and

the Effect of Prejudice on the Avoidance of Outgroup Friends. *Social Psychology Quarterly.* 78 (2). 127-50.

Tajfel, H., and Turner, J. C. (1979). An integrative theory of intergroup conflict. *The social psychology of intergroup relations*, 33, 47.

Turner, JC, R J Brown, and H Tajfel. (1979). Social Comparison and Group Interest in in-group Favouritism. *European Journal of Social Psychology* 9(2). 187–204.

Weiß, R. H. (2006). *Grundintelligenztest Skala 2 Revision (CFT–20R).* Göttingen, Germany: Hogrefe.

Rusbult, C. E. and Buunk, B. P. (1993). Commitment processes in close relationships: An interdependence analysis. *Journal of Social and Personal Relationships.* 10(2). 175-204.

Tajfel, H. and Turner, J. C. (1979). An integrative theory of intergroup conflict. *The social psychology of intergroup relations.* 33-47.

Vaquera, E. and Kao. G. (2008). Do you like me as much as I like you? Friendship reciprocity and its effects on school outcomes among adolescents. *Social science research. 37(1). 55-72.*

Wittek, M., Kroneberg, C. and Lämmermann, K. (2019) Who is fighting with whom? How ethnic origin shapes friendship, dislike, and physical violence relations in German secondary schools. Social Networks.

4 THE STRUCTURES WE LIVE IN KEEP US WITH OUR OWN? IN-GROUP TENDENCIES FOR ENDOGENOUS NETWORK PROCESSES AND THEIR ROLE FOR SEGREGATION

4.1 Introduction

At least, since the publication of the work by McPherson et al. (2001) the concept of homophily has become well known and received increasing attention by research that investigates for example adolescents inter-ethnic relations (e.g. Jugert et al. 2018; Leszczensky and Pink 2015; Boda and Néray 2015). Homophily refers to the idea that individuals are more likely to form and keep social relations and bonds with others who are similar in at least one characteristic, such as gender, race, social class or age[26]. With the increasing possibility to collect and analyse social network data, a growing amount of research found evidence confirming homophily in pupils' friendship networks where ethnicity tends to be a strong predictor for ties to be predominantly formed between members of the same ethnic group. With it, comes the danger of implicitly assuming that observing the macro-level state of a network to show more relationships between individuals with the same characteristics is the result of the micro-level mechanisms of choice homophily, understood as an *in-group preference* (Moody 2001; Baerveldt et al., 2004; Hallinan and Williams, 1989; McPherson et al, 2001; Smith et al., 2014; Verbrugge, 1977; Vermeij et al., 2009).

Drawing this conclusion could be misleading, because of well-known endogenous network mechanisms of social networks such as (but not limited to) transitivity (friends of friends becoming friends) (i.e. Goodreau et al. 2009, Feld and Elmore 1982), reciprocity (the tendency of friendships to be mutual rather than one-sided) and

26 Or combinations of multiple dimensions of homophily; e.g. Block and Grund 2014

centrality. Studies with a research focus on possible relations between homophily and endogenous network processes (Goodreau et al., 2009; Grund (2014); Grund and Densley, 2015; Kossinets and Watts, 2009; Wimmer and Lewis, 2010) clarify the importance of accounting for the latter when the research interest is focused on one of the others.

However, it is not always straight forward to disentangle network mechanisms from individuals' choices about whom to be connected to. Some network structures might show a tendency to arise predominantly within groups, because they reflect social situations that alter individuals' choices for rather having in- than out-group relations in the respective constellation. Such mechanisms are here referred to as structural homophily. From this derives the understanding of choice homophily in this work as being approached by accounting for homophily in endogenous network structures.

When investigating cross-sectional network observations instead of studying the longitudinal evolution of networks over time, it is especially difficult to differentiate both mechanisms. While it is possible in a cross-sectional observation to analyse if there are patterns of specific network structures to rather occur within or between groups, there could be different network processes leading to the same outcomes.

Thus, the primary goal of this study is to provide an example of how to get a better understanding about the role that choice homophily and structural homophily on the micro-level can play for the emergence and dynamics of segregation on the macro-level of a network. To this end, the endogenous network effects of reciprocity, transitivity and centrality are theoretically linked to the example of ethnic homophily of pupils in classroom friendship networks. In this study, a longitudinal perspective is applied which investigates the changes that pupils make within their friendship network, in order to better differentiate if certain network structures have a tendency to emerge within rather

then between groups. This is done in two steps, First, a sample of classroom friendship networks is analysed by stochastic actor oriented models (SAOM) to assess the presence of choice and structural homophily. These analyses will be used to evaluate the role of accounting for structural homophily in order to assess dynamics of segregation, Secondly, the estimated parameters will be used in a data-driven agent based simulations to investigate how structural homophily contributes to the emergence of ethnic segregation in classroom friendship networks. This data driven simulation approach can be understood to try to explain which homophily mechanisms might have been present during the emergence of the network leading to the initial segregation of the first observation.

4.2 Theory

4.2.1 Segregation and homophily – macro-state and micro-process

In contrast to existing work (e.g. Goodreau et al. 2009), here segregation is differentiated as a macro-level characteristic describing the state of a network at a given timepoint *t*, whereas homophily is understood as a micro-level process; actors in a network making systematic choices of creating and maintaining ties within groups rather than between. Regarding networks, this understanding of segregation refers to describing a state of 'inequality in the distribution of people across groups' (Rodrigues-Moral and Vorsatz 2016). While this is consistent with respect to connected actors being unequally distributed to be from the same group rather than from different groups, it can be misleading in cases where relations between actors are conditional on their group association (e.g. tutor and tutee are determining advice relation between students). While this work refers to 'group' categories as constant ascribed characteristics of individuals such as gender or ethnicity, this could also refer to categories that vary and are influenced by existing social relations (e.g. musical taste). Summarising, segregation

is understood as a characteristic of a network (as a whole) that describes the tendency of actors from the same group to have ties to one another rather than to those of another group, such as friendships in a classroom being more common between boys and boys and girls and girls rather than across sexes.

Homophily, on the other hand, is understood here as a mechanism on the level of individuals in networks making choices about whom they do or do not want to be related to. Simply put, homophily can be observed when individuals choose to be connected to others from their own groups. Pupils making choices about whom they want to be friends with is not necessarily a fully conscious decision comparing all possible options and weighing in all pro's or con's (e.g. the alter having the same migration background or not). Relating to social exchange theory (Emerson 1976; Rusbult and Buunk 1993), it is understood that pupils being friends with one another is the outcome of repeated interactions between the involved actors. If one pupil nominates another as their friend, this is a declaration that the quality of their relation reached a sort of threshold. The quality of the interpersonal relation improves from interactions that are perceived as personally beneficial for the involved individuals which could be by receiving resources, establishing personal or emotional commitments and trust.

Individuals will commit different types of investment into the other, such as time, sharing personal information or spending money on leisure time activities. In order to initiate and establish such interactions, individuals have to overcome and re-evaluate *interpersonal* uncertainties about whether their investments into the relation will pay off. For example a student will consider whether it is worth spending some pocket money on going to a specific movie with another student, or whether one can share a personal opinion about something without being made fun of.

In this process, homophily results from individuals making different assessments of their investments into relations with alters from a different group than with alters from the same group. When interacting with out-group alters, individuals face additional *intergroup* uncertainties or costs about the their investments into the relation compared to in-group alters. These can arise from different sources depending on the groups both are affiliated with. Following group identity theory (Taijfel and Turner, 1979), ethnic groups are one of the relevant characteristics that adolescents tend to identify with. For groups with different ethnic or migration background, these *intergroup* uncertainties that influence friendship choices could stem from existing prejudice, negative out-group attitudes (e.g. Schachner et al. 2015) or based on egos' expectations about different cultural norms and values. Under these assumptions, relationships in dyads of actors from the same group hold lower perceived risk of investment than between actors of different groups, because the first faces only *interpersonal* and the latter also *intergroup* uncertainties. In this framework, pupils show a preference to name others as their friends if they are from the same group rather than from another, is a resultof the first being perceived as a low-cost scenario when compared to the second.

This definition of homophily can be seen as more closely referring to the understanding of Lazarsfeld and Merton (1954) that actors' preferences are (part of) a process that can lead to segregation on the group level. However, in this work homophily is not necessarily ascribed to an in-group preference of students rather wanting to be friends with their own and more closely relates to baseline homophily (McPherson et al. 2001). Direct measures of such preferences are difficult to collect and thus are rarely available for empirical studies.

Deriving these preferences from the choices individuals make within networks of interdependently connected individuals is not straight forward. This is because the

observation of an existing relation between two actors can be (partly) the result of other relations these actors already have. Therefore, the goal of this work is to disentangle the respective network structures in which friendship ties to the in- or out-group are embedded to get a better understanding of the driving forces leading to segregated networks.

While segregation can be measured for any observed network at a given time *t*, by relating observed ties within groups and between groups, homophily referring to a process cannot easily be derived from a cross-sectional observation. This differentiation might not be necessary when observations of individuals or actors are independent, because it is possible to calculate a baseline of expected "segregation" if observed actors make choices disregarding respective group membership. Under these assumptions, the (cross-sectional) observation of a (de-)segregated network (one which deviates from the calculated baseline) can be attributed to actors' choices being influenced by group membership, if alternative mechanisms are accounted for (.e.g socioeconomic status). However, this must not be true for the case of interdependent observations, as is the case in many social networks such as friendships. Observing a friendship nomination from pupil A to pupil B must not only be the result of the respective characteristics of these two (such as their ethnicity) but can also be caused by the presence or absence of other ties in the network, such as an existing nomination from pupil B to A (*reciprocity*). These interdependencies between observations of ties are here described as *endogeneous network mechanisms.* In the cross-section of a network one can investigate to what extent network structures associated with these network mechanisms are present (e.g. how many reciprocal dyads exist), while a longitudinal perspective can investigate the emergence and stability of these structures.

A longitudinal approach not only allows for the possibility to investigate scenarios that

can lead to observed network states, but also the dynamics of the constituting structures. When specific network structures have different tendencies for change depending on whether they exist within or between groups, this can have a strong impact on the emergence and dynamics of segregation in a network. Generally speaking, tendencies for emergence and stability of these network structures favoring in-group members (*homophily*) would be expected to increase segregation. The next section briefly introduces the endogeneous network processes associated with reciprocity, transitivity and indegree centrality. Furthermore, it will be argued that these network processes can reflect different social situations and thus play different roles when relating to in-group or out-group dyads (i.e. ties between pairs of actors with the same or different migration background respectively). The following section provides arguments in line with rational choice theory about which of the micro-level network processes are expected to show tendencies for homophily and therefore might play a stronger role for the emergence and change of macro-level segregation.

4.2.2 Reciprocity

Reciprocated friendships provide benefits for both pupils involved, such as mutual social confirmation, trust and social capital (e.g. Rivera et al. 2010). Observing one-sided or asymmetric friendship nominations on the other hand can be a signal of status differences, personal strain or cognitive dissonance. Especially with regard to friendships it is therefore not unexpected that several studies find strong tendencies for reciprocity among youths in schools (e.g. Jugert et al. 2018; Leszczensky and Pink 2015; Boda and Néray 2015). As outlined above and in line with the reasoning in Block (2015), dyads within groups come at lower costs than between groups. Since reciprocal friendships provide more benefits than asymmetric ones, reciprocity can be expected to be less likely within than between groups. In other words, reciprocation within groups is

not as important for actors to create and maintain friendships than between groups. This found support by Hellpap et al. (chapter 3), where homophily in creation and maintenance was found only for one-sided friendships. Moreover, a strong tendency of heterophily was found for the maintenance of reciprocrated friendships, indicating that reciprocated friendships are more likely to persist between groups than within. Thus, when not accounting for homophily in asymmetric and reciprocal friendships separately one may underestimate the emergence of one-sided in-group ties over out-group ties and the maintenance over reciprocal out-group ties over in-group ties. However, these effects showed to differ considerably between classrooms, and this variance could not be related to their ethnic composition.

4.2.3 Transitivity

Transitivity refers to the tendency for friends-of-friends becoming friends. This tendency can be explained by the mutual friend an ego has with an alter to provide trust and reduce uncertainties about investing into a friendship. Grund (2014) shows how transitive closure can amplify initial tendencies for homophily related to network size.

Without going into too much detail, this can be clarified by considering how a network evolves from an empty network at *t0* with a slight preference of actors to create ties to others from the own group. With each tie created between actors of the same group over ties created between actors of different groups, the chances increase to be indirectly linked with someone from the own group rather than the out-group. Even if the tendency of transitive closure within groups is as large as between groups, numerically more ties will be created within groups thereby contributing to segregation. Thus, a stronger general tendency for transitivity can amplify small preferences for in-group tie creation to lead to strong segregation on the macro level. It follows that this amplification would be even stronger in the presence of homophily for transitive

closure. However, in line with the reasoning of Block (2015) that relationship within groups provide lower costs than relationships between groups, (transitive) triplets formed by three individuals can be understood as (small) groups. Therefore, they provide a lower cost situation for individuals than non-embedded dyadic relations. Since inter-group relations constitute high cost situations compared to intra-group relations, the first might benefit stronger from being embedded through an intermediately connected actor than the latter. Analogue to the argumentation made for reciprocity, transitivity might play a stronger role for the formation and maintenance of inter-group friendships. Therefore, homophily should be higher in dyads that are not embedded through inter-connected actors compared to those that are. This suggests that homophily in transitive triplets should be low if present at all, or even show small tendencies for heterophily in transitivity transitive ties being more prominent between than within groups. On the other hand, the presence of pronounced group boundaries and individuals' strongly identifying with their own ethnic group, could result in them seeking to enforce ethnic boundaries by confirming group identification through intra-group transitivity. Either way, not accounting for homophily (or heterophily) in transitive triplets if it is present would lead to underestimating (or overestimating) homophily in not-embedded dyads and thus predict lower (higher) levels of segregation.

4.2.4 In-degree centrality

The third endogenous network process relates to central actors. While there are several ways centrality can be conceptualized (e.g. Borgatti, 2005; Opsahl et al. 2010), this chapter looks only at in-degree centrality. In our context, central actors are those pupils that receive a high number of ingoing friendship nominations. These high indegree nodes can play an important role in the process of segregation. For example, their social actions can have an influence on those that associate with them. Thus, if they have a

preference for befriending others from the in- or out-group, this behavior might be copied and increase or respectively decrease segregation. In this work however, the focus is on the role of the structural network mechanisms.

Central actors can be seen as facilitators for triadic closure. In other words, they are pupils that interconnect several others who then are friends of friends. As outlined earlier, a tendency for triadic closure can amplify low levels of homophily and so lead to segregated networks. Similarly, I argue that indegree centrality plays a role in the process of network segregation since it provides opportunity for triadic closure. Pupils that are central within a group establish more open two-paths between pupils of the same group. This increases segregation when triadic closure occurs. On the other hand, central pupils that are similarly connected within and between groups can counter-act segregation in friendship networks.

Of course, when interested in the emergence of segregation and associated network structures, one hast to consider how actors become more or less central within or between groups. Iit can be argued that pupils have a tendency against creating or maintaining ties to those who already have several ingoing nominations. One possible explanation for this is that central actors do not reciprocate all their ingoing ties. Students who are not nominated back by these central actors might not maintain their ties to them, which would result in these actors becoming less central. If non-reciprocated ties rather persist within groups, it is likely that actors become more central within their groups rather than connecting between them. Thus, central actors becoming more or less popular within or between groups can have a strong impact on the emergence and dynamic of segregation.

Reconsidering the presented theoretical arguments underlines the importance of considering mechanisms behind the emergence and evolution of networks to become

more or less segregated. The endogenous network mechanisms can show different tendencies of homophily, while also clearing the baseline homophily of it's structural network components. Secondly, taking homophily in endogenous network processes into account provides a more flexible model of different potentially inter-dependent forces that lead to more or less segregated networks. Therefore, it is expected that accounting for in-group tendencies of the three endogenous network mechanisms is better able to simulate the emergence (**H1**) and change (**H2**) of segregation than accounting only for dyadic homophily.

4.3 Data and Method

This study uses 64 classrooms from the first two waves of the Swedish CILS4EU dataset (Kalter 2016). This exemplary subsample is selected from the 251 classrooms available in the dataset based on suitability for model fitting. Since all analysed classrooms need to be fitted with the same set of parameters, only those remain in the sample for which the different models produce converging results. The main cause for excluding networks from the analysis is mainly that the models overspecify homophily which leads to collinearity between parameters of interest. This is most likely to occur in two types of classrooms. On the one hand, classes that have an unfavourable combination of the ethnic composition in the classroom with one group constituting a large majority of the pupils and very few relationships between groups. These classes do not provide enough information to differentiate the the three types of structural homophily parameters or for some instances even choice homophily. The second set of classes that did not result in converging results are those where homophily (or heterophily) mainly occurs in one of the endogenous network processes. For example this would the case if nearly all inter-group friendships that are created or maintained are reciprocal and intra-group friendships being mainly one-sided. Both could imply

that the sample selection provides a bias towards the mean for the homophily specifications of interest, because it excludes cases that are strongly driven by at least one of the mechanisms. While such a selection is obviously not suited to be generalisable to larger population (e.g. of friendship networks in Swedish secondary school classes), this is not necessarily detrimental to the goal of this paper which is to understand the role of choice and structural homophily for network segregation.

First, parsimonious stochastic actor oriented models (Snijders et al. 2010) are fitted separately to the networks for five different models using the software package RSiena (Ripley et al. 2019). For the purpose of this study the framework of actor oriented models is the preferred approach because it is more congruent with the motivation to investigate how network structures can affect individual's choices for network changes. Applying the hierarchical stochastic actor oriented models (Koskinen and Snijders 2019) would promise to produce estimates for a larger proportion of the sample, because of the Bayesian estimation procedure. However, the equivalent method for assessing goodness of fit which is utilised here is to assess the role on structural homophily for segregation dynamics is still being developed for the HSAOM. Thus, single network SAOMs have been estimated for the set of selected model specifications. The first baseline model accounts for gender homophily, and the three baseline network processes; reciprocity, transitivity (measured as geometrically weighted shared partners GWESP[27]) and indegree popularity (**M1**). The second model adds a basic dyadic effect for ethnic homophily, which accounts for a general tendency of actors creating and maintaining ties within rather than between groups (**M2**). Thirdly, one model is estimated including an effect that additionally accounts for in-group tendencies in each of the basic network processes seperately (**M3a**-**M3c**), while the last model includes the

27 The default tuning parameter of $\alpha = .69$ was used in accordance with Snijders et al. (2006). As is the case for ERGMs, this parameter indicates the decay or diminishing returns of each additional intermediate actors contributing to the creation or maintenance of the respective transitive tie.

effects for homophily in network mechanisms simultaneously (**M3**). All models also include a density effect which represents actors' average outdegree net of all other parameters, as well as a rate parameter representing the average changes made by actors between observations. Each estimation has been run with default of 1000 iterations of the main-phase estimation, resulting in 1000 simulated networks, given the specified model. In order to test if accounting for structural homophily leads to better assessment of the segregation dynamics (**H2**) the difference in segregation between simulated networks and observed networks is compared between **M2** and **M3**. This approach follows the logic of assessing the goodness of fit for different models in order to link micro-processes with macro-outcomes (Snijders and Steglich, 2015).

Following the procedure in Stadtfeld et al (2018b), in the next step each set of estimated parameters is used to simulate 100 chains of up to 500 times the estimated rate for each of the respective networks. The simulations are carried out as stochastic actor oriented models starting from an empty network with all actors present at the first observation and can be understood as agent based models (Snijders and Steglich, 2015; Prell and Lo, 2016; Block et al., 2018; Stadtfeld et al., 2018a). Hereby, the rate parameters represent the average changes actors make and can be understood as time in our simulation model. To illustrate this, an estimated rate parameter of 5 means that in average pupils in the respective classroom make 5 changes in their network between the observations. Simulating chains of 500 times the rate means that the network evolution is simulated from an empty network to a point where on average, pupils made 2500 changes in their network.

In order to investigate the dynamics of the simulated emerging networks, segregation levels are calculated for each classroom after each subsequent rate in the simulated chains and averaged over the 100 chains. This allows studying segregation dynamics at

the early stages of the network evolution as well as trends towards an equilibrium in the long run. Results are presented for the averages over all classrooms for each of the models. However, the classrooms might differ from the overall tendencies of segregation. Therefore, in a second step the segregation values from the chains of simulated networks will be compared to the segregation levels of the first network observation for each of the classroom. Investigating this type of bias measure should help to compare which of the models produces simulations that are better in grasping the emergence of segregation. While the first step analyses homophily in endogenous network effects between the two observations and their contribution to explaining dynamics of segregation, the second step investigates if these are valid mechanisms to explain the emergence of observed segregation levels. When comparing the data-driven agent based simulations to the first wave of the network observation, we assess if the segregation in this network is the result of the dynamics modelled between both observations. In other words, we simulate a possible process of an emerging network leading to the segregation at our first observation under the assumption that the network emerged from the mechanisms specified in the estimation being constant over time. Of course this strong assumption about network mechanisms being homogeneous in time over the process of network emergence would be very rarely plausible. However, here it serves as a simplified example to use the method introduced by (Stadtfeld et al 2018b) for this purpose. The simplified models that are applied and the assumption about time homogeneity should be relaxed and thoroughly investigated.

4.4 Results

A first look on the descriptive statistics in table 4.1 shows for the analytical sample that the 64 classrooms have between 16 and 34 averaging about 25 pupils. The classrooms consist of an average of 51.8 percent females ranging from 21 to 87 percent. The

average proportion of non-native pupils is 29.4 percent, ranging from 3.6 to 94.1 percent. This implies that some of the classrooms have very few native or non-native pupils which can restrict the opportunity structure for in- or out-group ties of the respective groups. This would be especially problematic when investigating group specific differences of in- and out-group ties, which is not the case here. Thus, these classes have been kept in the analyses to not further restrict the sample.

The density of the friendship networks has a mean of 13.7 percent realised ties in the first wave which is slightly reduced to 12.4 percent in the second wave. This reduction reflects in pupils nominating on average 3.2 others as their best friends in wave 1 and 2.8 at the second wave. As can be expected, dyadic reciprocity is quite high, revealing that about 89.8 percent of the dyads with ties are reciprocal, which reduces marginally to 89.1 percent at wave 2. This illustrates the importance of mutual nominations for the friendship network. The clustering coefficient, or transitivity, indicates that the probability of two pupils who have a mutual friend are friends with each other is initially .561 and reduces slightly to .531 during the second observation. The indegree centrality scores indicate only small tendencies for centralized network structures with an average of .167 at wave 1 and .169 at wave 2. While the networks are strongly reciprocal, clustered but only slightly centralized, the reduction of friendship nominations might rather result from transitivity going down.

Next, the network segregation is investigated by looking at the the difference of in-group and between-group densities. The measure is calculated by subtracting the proportion of realised ties from possible ties between actors of different groups from the proportion of realised ties from all possible ties between actors of the same group. This measure could be understood as an extension of the EI-Index that not only takes into account differences of existing ties in a network, but also differences in absent ties. The

minimum value of the measure is -1 indicates completely desegregated networks where all possible ties have been realised between actor of different groups but no ties are present between actors of the same group. The maximum value of +1 represents the opposite case of completely segregated networks were all ties between actors of the same group are realised but no values between actors of different groups. While this measure would not be ideal for assessing segregation based on continues node characteristics, it provides a clearly interpretable scale for categorical characteristics. However, for this study it was chosen because it accounts for the opportunity structure, which is essential for comparisons over time and between networks of different size and with different compositions, Moreover, when calculating the possible ties within and between groups it is easily possible to account for the restriction of pupils not being able to nominate more than five alters as friends, which results from the design of the survey. This measure will be used in all following sections as indicator for the macro-level outcome of interest.

. It can be seen that within the classrooms at the first wave on average 30.4 percent more ties are realised within groups than between, which reduces to 25.3 percent in the subsequent wave. These cross-sectional descriptives give a first indication about the trends of the classrooms between the two observations, but they do not provide much analytical insights about the individual changes pupils make with their friendship choices.

Classroom level				
	min	mean	max	sd
N pupils	16	24.6	34	3.8
% non-native	3.6	29.4	94.1	21.3
% female	21.4	51.8	87.0	12.7

Friendship Network					Wave 2			
	min	mean	max	sd	min	mean	max	sd
Degree	1.947	3.172	4.231	.532	1.741	2.844	4.074	.527
Density	.087	.137	.196	.027	.061	.124	.213	.031
Reciprocity	.821	.898	.967	.030	.735	.891	.947	.040
Transitivity	.326	.561	.797	.109	.247	.531	.789	.105
Centrality	.064	.167	.325	.053	.087	.169	.301	.047
Segregation	.021	.304	.835	.183	-.042	.253	.781	.183

N classes = 64

Table 4.1: Basic descriptives

Tables 4.2 and 4.3 show the results of the conducted classroom-wise stochastic actor oriented models reported as a meta-analysis directed at testing for parameters being different from zero, based on Fisher's method for combining independent p-values (Fisher 1932; Hedges and Olkin 1985; Snijders and Bosker 2012). The analyses were conducted using the implemented function from the RSiena software package, which following Viechtbauer (2010) represents a random effect approach with a Hedges estimator (Ripley et al. 2019). This method was chosen, because the subset of networks that were analysed here can not justifiably be treated as sample from the population of dynamic friendship networks in Swedish school classes for which a mean effect and standard error is calculated. Since the goal of the study is not in estimating generalisable effect sizes, but to provide an example of how the role of structural homophily for the emergence and dynamic of segregations can be investigated, this method serves the purpose of providing initial information about the parameter distributions.

The baseline model in column **M1**, first shows a negative outdegree effect (b=-1.617,

se=.090, p<.001) reflecting that pupils nominate fewer friends in the second wave than the first. The model also shows tendency for gender homophily (b = .213, se=.075, p <.01). The endogenous network effects of interest, show a strong tendency for the creation and maintenance of reciprocal ties (b=1.452, se=.077, p <.001), transitive triplets (b=1.426, se=.0.068, p < .001) and a tendency against nominating pupils who already have more ingoing nominations (b= -.0.115, se= .020, p<.001). The next model in coloumn **M2** additionally includes an effect for basic homophily for the groups of native and non-native pupils which shows a positive tendency for in-group tie creation and maintenance (b=.270, se=.060, p<.001). Including this effect does not substantially affect the other parameters in the model. The last coloumn **M3** additionally accounts for the discussed in-group tendencies of the three endogeneous network effects. However, this model does not show any indication of homophily, neither for the general effect nor the endogeneous network mechanisms.

Effect names	M1		M2		M3	
	b	s.e.	b	s.e.	b	s.e.
Outdegree	-1.617	.090 ***	-1.793	.113 ***	-1.655	.124 ***
Gender homophily	.213	.075 **	.253	.079 **	.266	.080 **
Reciprocity	1.452	.077 ***	1.432	.079 ***	1.345	.106 ***
gwesp	1.426	.068 ***	1.410	.070 ***	1.471	.107 ***
Indegree popularity	-.115	.020 ***	-.116	.021 ***	-.128	.028 ***
Migrant background homophily			.270	.060 ***	.166	.113
X reciprocity					.138	.122
X gwesp					.004	.124
X indegree popularity					.020	.032

*** p < 0.001 ** p < .01 * p < .05

Table 4.2: Results of meta-analysis of SAOM in 64 classrooms

Results for the models that include the effects for in-group tendencies of the endogenous network effects separately are shown in table 4.3. None of the columns

M3a to **M3c** show evidence for in-group tendencies of any of the endogenous network processes reciprocity, transitivity or indegree popularity. However, in the model including a parameter for in-group tendencies of transitivity in **M3b** we find that the otherwise positive and significant effect of general dyadic homophily becomes non-significant.

Effect names	M3a		M3b		M3c	
	b	s.e.	b	s.e.	b	s.e.
Outdegree	-1.730	.120 ***	-1.738	.133 ***	-1.764	.120 ***
Gender homophily	.230	.073 **	.255	.080 **	.249	.077 **
Reciprocity	1.310	.110 ***	1.437	.082 ***	1.445	.080 ***
Transitivity	1.429	.073 ***	1.408	.107 ***	1.433	.076 ***
Indegree popularity	-.113	.021 ***	-.119	.022 ***	-.129	.022 ***
Migrant background homophily	.175	.072 **	.201	.105	.226	.065 ***
X reciprocity	.108	.119				
X transitivity			.029	.119		
X indegree popularity					.029	.030

*** p < 0.001 ** p < .01 * p < .05

Table 4.3: Results of meta-analysis of SAOM in 64 classrooms

The next step investigates how the modelled micro-level mechanisms of structural homophily contribute to the explanation of segregation on the macro-level. For this, an approach similar to the goodness of fit method as described in (Snijders et al 2010b and Snijders and Steglich 2015) is used. For this, the previously introduced measure of segregation based on the difference of in-group and between-group densities is calculated for each network. Following, for each network the 1000 simulated networks resulting from the iterations of the main-phase of the estimation procedure are extracted for M2 and M3 and their segregation level is calculated. Next, For each network the difference between the observed segregation in the network and the segregation in each of the 1000 simulated networks is calculated by subtracting the first from the latter, for

both models. The resulting values are positive if a simulated network shows more segregation than observed and negative values if the segregation of the simulated network is lower than in the observed network. These differences between the segregation from the observed and simulated networks are here referred to as bias. For each network and model, the average of these 1000 bias values is taken to indicate if the model specification accounting for none or all three structural homophily mechanisms tends to rather lead to more or less segregated network than the observed one. This leads to one average bias values per network and model. These average bias values are then compared between models by mean comparison. The average difference in bias between models is very close to zero ($\bar{x}$<.001, sd= .01) and varies between about -3% and 4% of more or less group ties being realised within than between groups, when comparing to the observed network. Applying t-tests on the differences in bias between the two models for each classrooms does not provide support for **H2**. While about a third (n=23) of the classrooms show significantly smaller bias when accounting for homophily in endogenous network mechanisms, about one third shows the opposite (n=24), with the remaining classes not showing differences between both models (n=17).

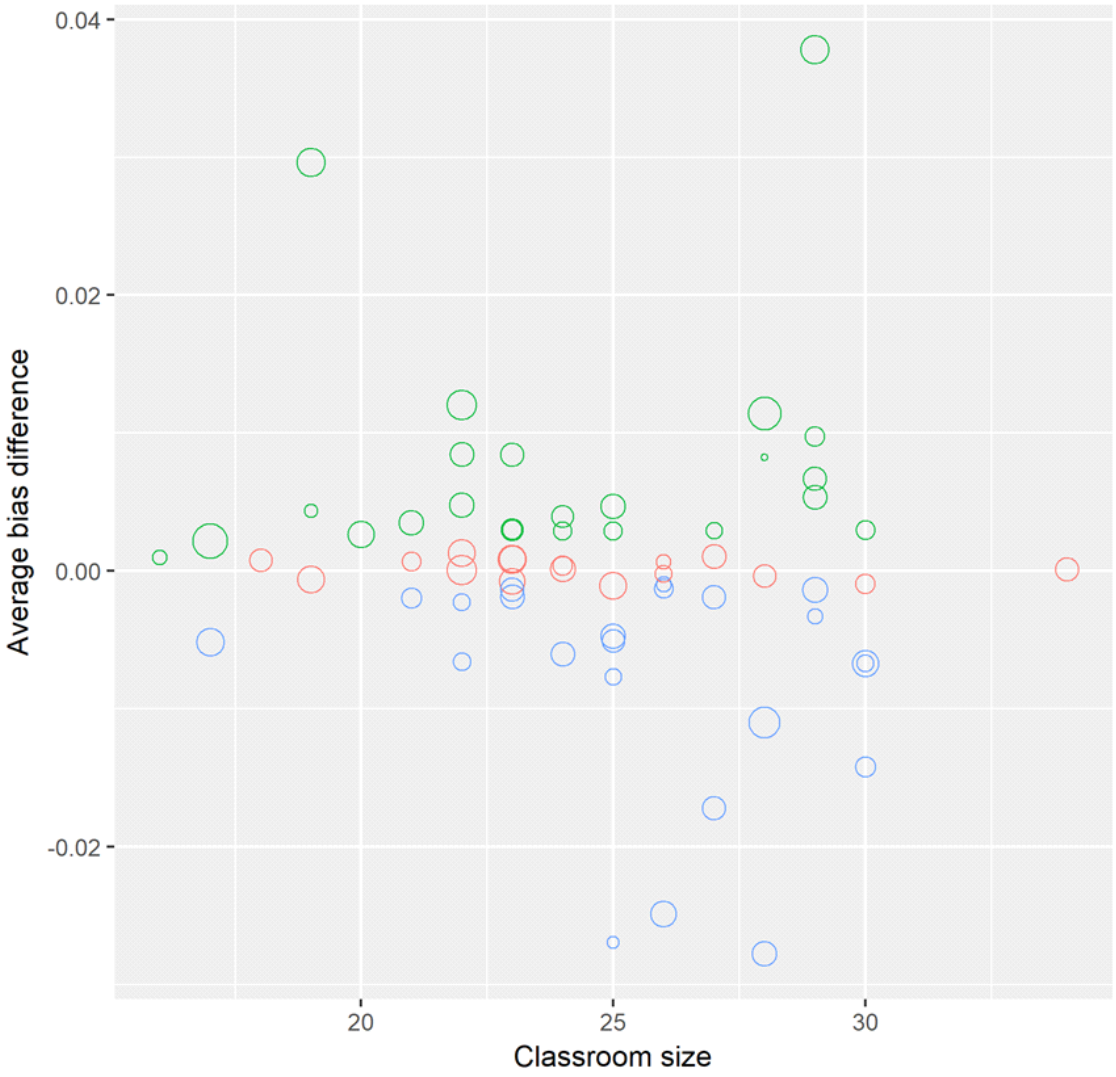


Figure 4.1: Average bias difference by classroom size

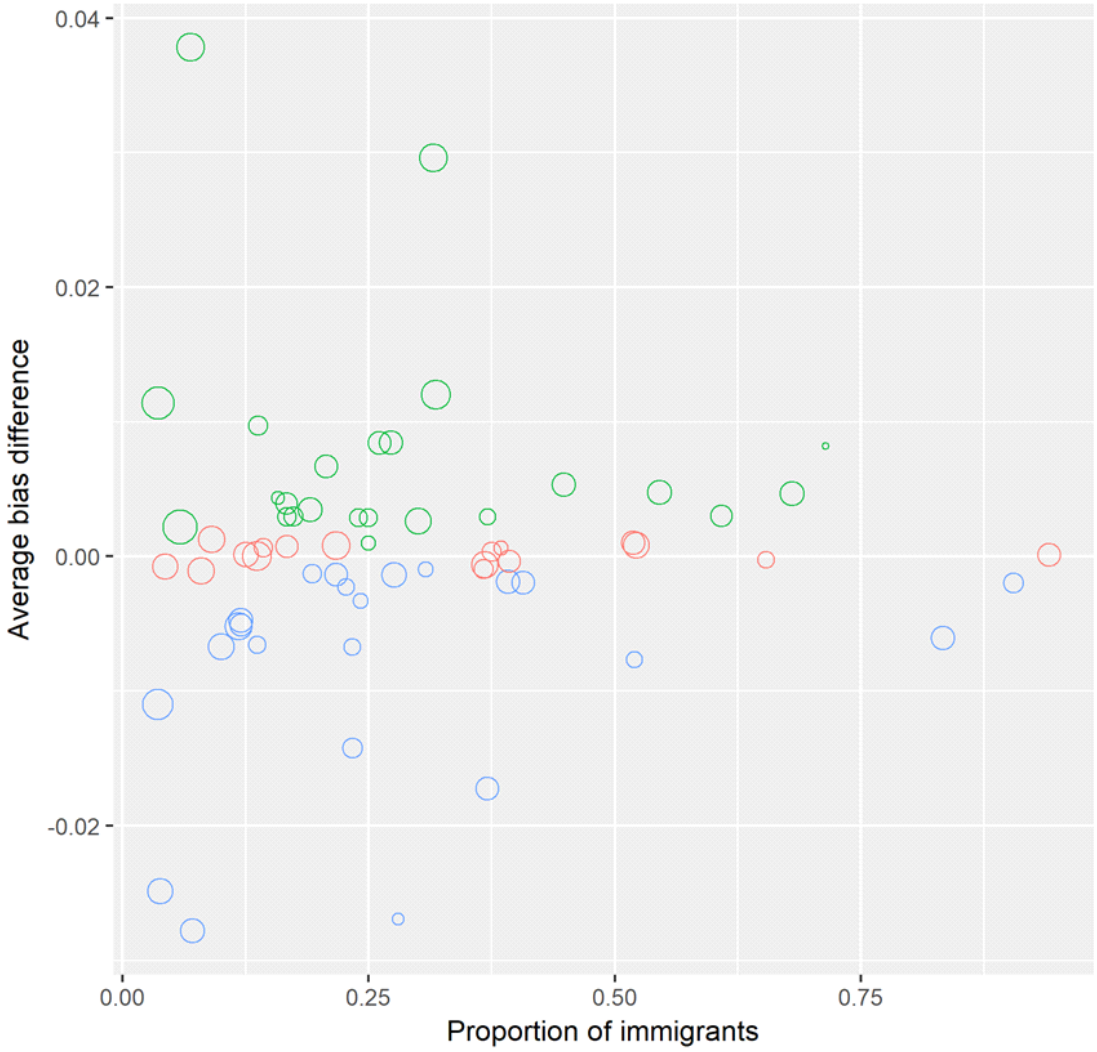


Figure 4.2: Average bias difference by proportion of immigrants

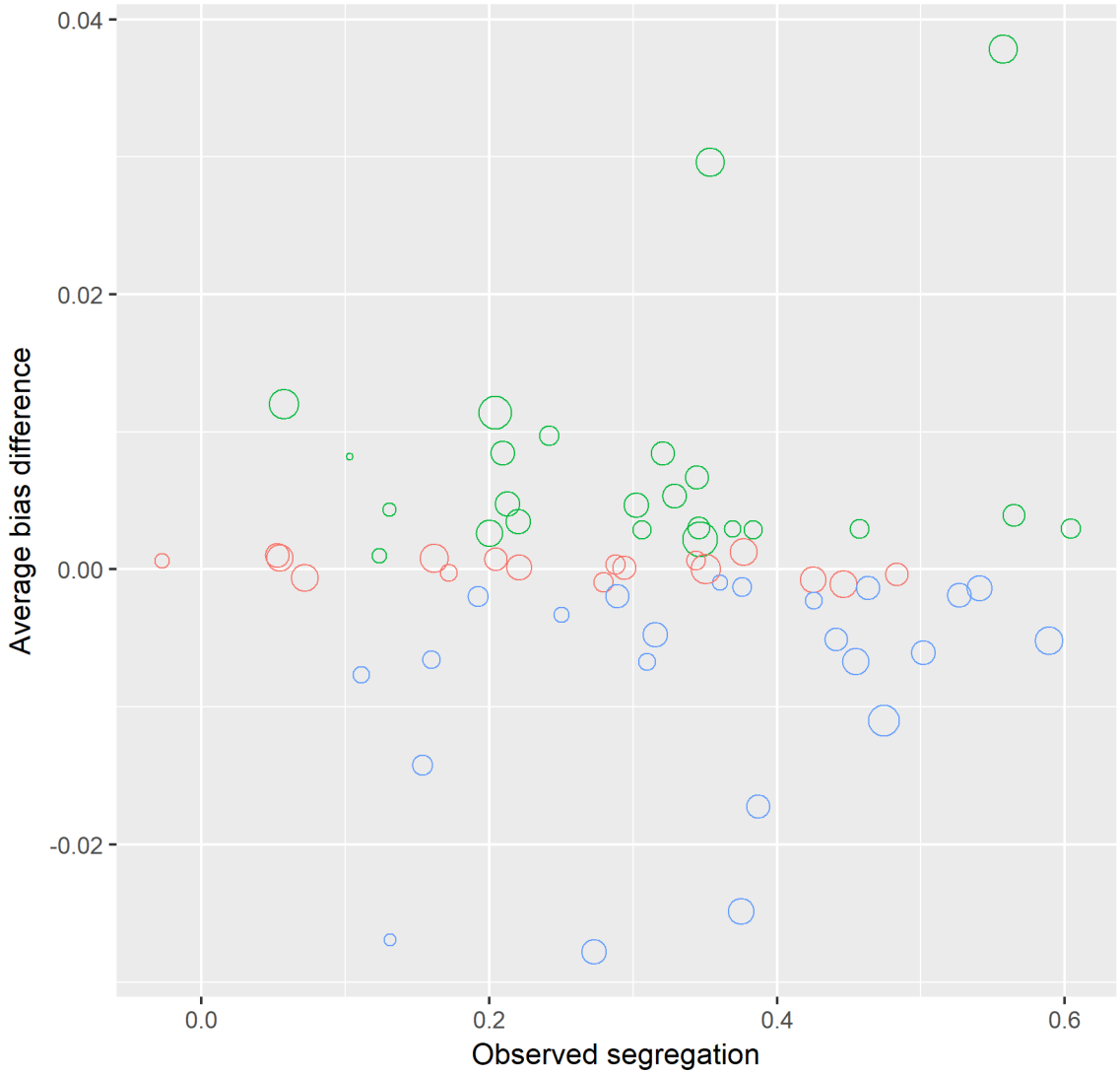


Figure 4.3: Average bias difference by observed segregation

Figures 4.1 to 4.3, show the difference in bias between the two models on the y-axis for each classroom by different classroom characteristics on the x-axis. Green circles mark classes where M3 produces simulated networks with segregation being significantly closer to the observed value, red circles indicate the opposite case, and black circles show those classes where differences between models are not significantly different. The size of the circles is equivalent to the standard deviation of bias between the simulated networks within each classroom. As can be seen, no clear trends links the differences to either classroom size, proportion of non-natives in the classroom nor

observed segregation during the second wave.

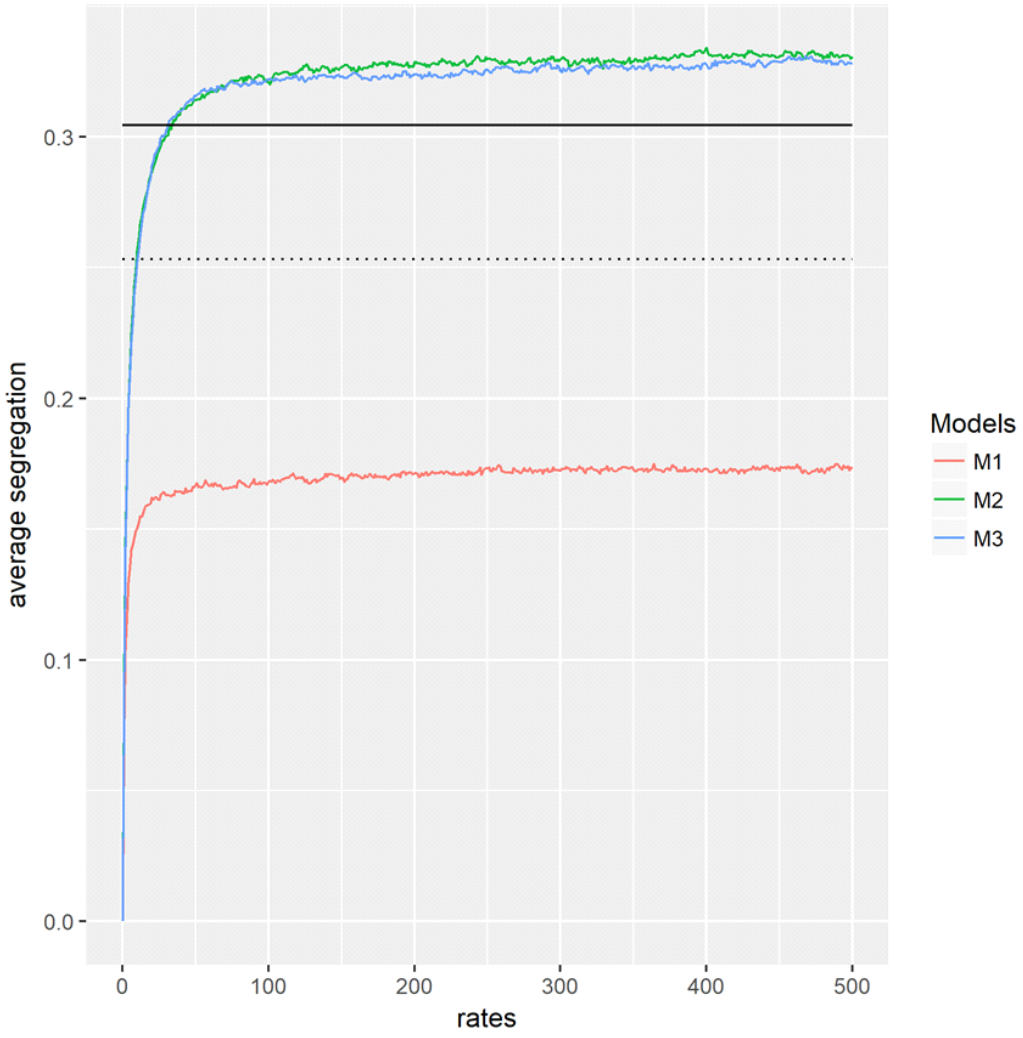


Figure 4.4: Segregation of agent-based simulation M1-M3

Figure 4.4. presents results from the agent-based simulations using the estimated classroom specific parameters of models M1-M3. The y-axis displays the overall mean segregation for all simulations of all classroom networks. The x-axis shows the model time in multiples of the estimated rate parameters. The horizontal lines represent the overall mean of segregation, with the solid line representing wave 1 and the dotted line wave 2. The different coloured lines indicate the estimated mean segregation. The first thing to be noted is that the segregation levels of the simulated networks tend towards

an equilibrium and seem to become rather stable within the first 100 observed rates. To no surprise, M1 including no parameter for homophily of immigrant background produces the lowest levels of segregation with the average staying at about 17 percent of realised in-group ties over between group ties. As can be seen, this model underestimates the observed segregation of wave 1 and wave 2 for all simulated networks. Including the basic homophily effect in M2 in fact produces the highest average of simulated segregation of around 33 percent. The model containing the three additional parameters of in-group tendencies in endogenous network mechanisms leads to simulations with slightly lower segregation levels of about 32 percent providing some support for **H1**. Both M2 and M3 reach higher levels of segregation than the observed value after the first 50 rates.

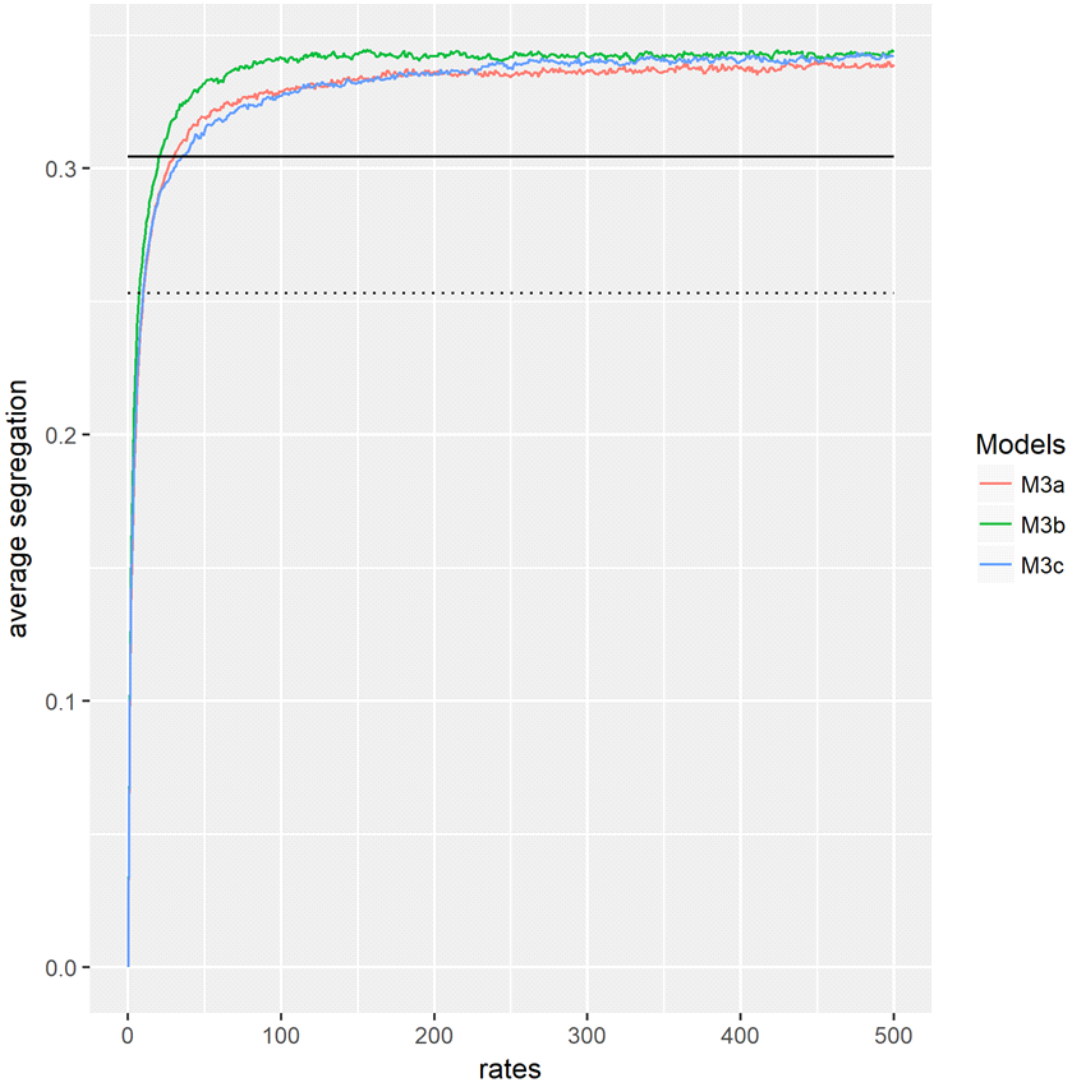


Figure 4.5: Segregation of agent-based simulation M3a-M3c

Figure 4.5. provides further insights about the role of network effects for ethnic segregation by showing the same measures but for the models m3a to m3c, which account for in-group tendencies of endogeneous network effects seperately. At a afirst glance, all three models show a convergence to a similar equilibrium at around 34% segregation producing higher than observed values before the first 50 rates. However, they show slightly different courses over the sequences of simulated network changes. Most notably, **M3b** the model accounting for tendencies of homophily in transitivity

seems to reach an equilibrium quickest with segregation levels rising the steepest at the beginning, followed by only small fluctuations for about half the remaining simulation time. Model **M3a** accounting for homophily in reciprocity has a slightly steeper rise than **M3c** but tends towards slightly lower values than the other two models. **M3c** including homophily in popularity has the lowest initial rise in segregation but continuous to increase slightly until the simulation reaches about 250 rates.

Lastly, the difference between the simulated segregation and the observed segregation at wave 1 is calculated for each classroom and each model after simulated rate. In order to test if one of the models produces simulated networks with segregation being significantly closer to the observed segregation, t-tests are run comparing all models with each other for every point of the simulated chains. However, none of the models shows to produce significantly better segregation values than any of the others, except when compared to the model without any parameters of ethnic homophily (M1) which produced significantly lower results of segregation than all other models.

4.5 Discussion and Conclusions

For a sample of 64 Swedish classrooms, this work provided an example of how to investigate the role of differentiating endogenous network mechanisms depending on whether they are situated within or between groups for the emergence of segregation in a network. While a meta-analysis of separate stochastic actor oriented models showed general homophily for immigrant background, in-group tendencies could not be found for neither reciprocity, nor transitivity or indegree popularity. The estimates of each classroom were then used as a basis for agent-based simulation in the stochastic actor oriented model framework, starting from an empty network and allowing for up to 500 times of the estimated average changes per student. The networks simulated while accounting for different structural homophily parameters can be compared to the first

observation of the network. This comparison can be understood as trying to simulate the process that resulted in the segregation of the network at the first observation under the assumption that the mechanisms modelled in the estimation are constant over time. Comparing the levels of segregation from the simulations between models revealed that including groupwise tendencies for endogenous network processes results in slightly lower levels of segregation in the simulated networks compared to including only a general dyadic homophily effect. While the baseline model without any parameters of ethnic homophily strongly underestimates the observed segregation, accounting for dyadic homophily only overestimates observed segregation even more than when including homophily for endogenous network mechanisms.

Endogeneous network mechanisms represent different social situations in which individuals are embedded and which have an impact on the changes they make on their surrounding network. This work outlines how the friendships that pupils establish with each other in their classroom can have an impact on whom they call their friends at later times and how this can result in more or less segregated networks based on their immigrant background. Over all classrooms, no significant evidence could be found that pupils rather reciprocate friendships within their own or to the other group. However, accounting for these tendencies in each classroom still helped to slightly better simulate the emergence of segregation compared to accounting only for basic dyadic homophily. Similarly this is also the case for pupils tendency to rather be friends with popular actors from their own or the other group. While the meta-analysis did not reveal a significant interaction term, the simulations produce average segregation levels that are closer to the observed segregation than the baseline homophily model.

Regarding pupils tendencies to befriend friends of their friends, no trend was revealed regarding in-group or out-group closure. Still, when including this mechanism in the

estimation the dyadic homophily effect turned insignificant. While it is possible that the networks are overfitted with this homophily interaction, this is not the case in the other models including these interactions where the dyadic homophily parameter remains significant. It might therefore be the case that a larger proportion of in-group dyads is created and maintained in embedded structures. The simulations from this model lead to the quickest rise in segregation, which speaks for the acceleration argument of transitivity with some in-group tendencies.

Not finding significant tendencies for homophily in endogeneous network mechanisms could be attributed to a lack of power, for this case of considerably small networks such as the analysed classrooms. This might pose a problem especially for those classes which have a very small proportion of native or non-native pupils. Since the dependent variable in the models refers to changes in the network, these classes might hold only few ties between groups. While some change in these ties might be sufficient to estimate baseline homophily, specifying further subsets of these occasions into the respective network structures could lead to small cell counts and thus lead to larger standard errors. However, the main goal of this study was not to reveal homophily in network mechanisms, but rather to investigate how such mechanisms can affect the emergence and evolution of a network and it's tendencies to segregation.

It can be argued that including additional homophily parameters in the simulation model leads to an increase of segregation in the resulting networks. This could be the case when the homophily interactions are positive, since they would generate and maintain more ties within groups during the simulated network evolution. However, this must not necessarily be the case, given that including these interactions also affects the dyadic baseline segregation to explicitly refer to not being embedded in the accounted network structures. As the findings illustrate, the homophily in endogenous network mechanism

models actually lead to lower levels of segregation than including only dyadic homophily and thus being closer to the observed segregation. This underlines that dyadic homophily in itself is not a sufficient mechanism to explain segregation in emerging networks. Still, in all cases the simulations based on ethnic homophily lead to overestimating the observed segregation and thus leaves open the question about meaningful mechanisms that help to improve simulation models. Future research building on this work, should first relax the assumption of mechanisms being constant over time. One of the first steps could be to differentiate mechanisms of tie creation and maintenance, since the first should play a stronger role in the early stages of the emerging network.

The approach applied here promises to be a useful tool in order to differentiate mechanisms that took place in the emergence of the initial network observation with regard to segregation levels. However, it might also be the case that the models applied here left out other relevant mechanisms that are at the core of the segregation dynamics. Different meaningful extensions can add to achieve this. First, Stadtfeld et al (2018b) from which the applied data driven simulation was adapted here, introduce this analytical method to show the relevance of the interplay between friendship and dislike networks to simulate formation of stable groups. The approach from Block and Grund (2014) points towards accounting for multidimensional homophily, who show that ethnic and gender homophily can be interrelated. While the here presented work can be extended by increasing the scope of networks of covariates, it demonstrated the relevance of all-present network mechanisms and the role they can play for pupils' network choices depending on whether they occur within or between groups.

4.6 Literature

Block, P. (2015). Reciprocity, transitivity, and the mysterious three-cycle. *Social Networks* 40. 163–73.

Block, P., and T. Grund. (2014). Multidimensional homophily in friendship networks. *Network Science.* 2 (2). 189–212.

Block, P., J. Koskinen, J. Hollway, C. Steglich, and C. Stadtfeld. (2018). Change we can believe in. Comparing Longitudinal Network Models on Consistency, Interpretability and Predictive Power. *Social Networks.* 52. 180–91.

Boda, Z. and B. Néray. (2015). Inter-ethnic friendship and negative ties in secondary school. *Social Networks.*43. 57-72.

Borgatti, S.P. (2005). Centrality and network flow. *Social Networks.*27(1). 55-71

Emerson, R. M. (1976). Social exchange theory. *Annual review of sociology*, 335-62.

Fisher, R. A. (1932). Statistical Methods for Research Workers. Oliver & Boyd, 4th edition

Goodreau, S. M, Kitts, J. A., Morris, Martina (2009) Birds of a Feather, or Friend of a Friend? Using Exponential Random Graph Models to Investigate Adolescent Social Networks. In: *Demography* 46 (1), 103–125.

Grund, T. (2014) Network Size and Homophily: Same-Sex Friendships in 595 Scandinavian Schools, in Manzo, G. (eds.) (2014) *Analytical Sociology. Norms, Actions, Networks*. New York: Wiley. 299-316.

Grund, T. U.; Densley, J. A. (2015) Ethnic Homophily and Triad Closure: Mapping Internal Gang Structure Using Exponential Random Graph Models. In: *Journal of Contemporary Criminal Justice* 31 (3), 354-70.

Hedges, L. V. and Olkin, I. (1985). Statistical Methods for Meta-analysis. New York: Academic Press.

Jugert, P., Leszczensky, L. and S. Pink. (2018). The Effects of Ethnic Minority Adolescents' Ethnic Self-Identification on Friendship Selection. *Journal of Research on Adolescence.* 28(2). 379-95.

Kalter, F., Heath, A.F., Hewstone, M. Jonsson, J.J., Kalmijn, M. Kogan. I., and F. van Tubergen. (2016) Children of Immigrants Longitudinal Survey in Four European Countries (CILS4EU) – Full version. Data file for on-site use. GESIS Data Archive, Cologne, ZA5353 Data file Version 1.2.0, doi:10.4232/cils4eu.5353.1.2.0.

Kalter, F., Heath, A.F., Hewstone, M. Jonsson, J.J., Kalmijn, M. Kogan. I., and F. van Tubergen. (2016) Children of Immigrants Longitudinal Survey in Four European Countries (CILS4EU) – Full version. Data file for on-site use. GESIS Data Archive, Cologne, ZA5353 Data file Version 1.2.0, doi:10.4232/cils4eu.5353.1.2.0.

Koskinen, J. H., Snijders, T. A. B. (2019): Multilevel longitudinal analysis of social networks (*in preparation*).

Kossinets, G., Watts, D. J. (2009): Origins of Homophily in an Evolving Social Network. In: *American Journal of Sociology* 115 (2), 405–50.

Lazarsfeld, Paul F.; Merton, Robert K. (1954): Friendship as a Social Process: A Substantive and Methodological Analysis. In: Morroe Berger, Theodore Abel and Charles H. Page (Eds.): Freedom and Control in Modern Society. Toronto: Van Nostrand, pp. 18–66.

Leszczensky, L., and S. Pink. (2015). Ethnic Segregation of Friendship Networks in School: Testing a Rational-Choice Argument of Differences in Ethnic Homophily between Classroom- and Grade-Level Networks. *Social Networks.* 42. 18-26.

McPherson, M., Smith-Lovin, L., and M.J. Cook. (2001). Birds of a feather: Homophily in social networks. *Annual review of sociology.* 415-44.

Opsahl, T., Agneessens, F. and J, Skvoretz. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks.* 32(3). 245-51.

Prell, C., and Y.-J. Lo. (2016). Network formation and knowledge gains. *The Journal of Mathematical Sociology.* 40.

Rivera, M. T., S. B. Soderstrom, and B. Uzzi. (2010). Dynamics of Dyads in Social Networks: Assortative, Relational, and Proximity Mechanisms. *Annual Review of Sociology.* 36(1). 91–115.

Rodrigues-Moral, A. and Vorsatz,M. (2016). An overview of the measurement of segregation: classical approaches and social network analysis. *Complex networks and dynamics, lecture notes in economics and mathematical systems* 683.

Schachner, M, K, Brenick, A., Noack P., Fons J.R. van de Vijver, and B. Heizmann. (2015). Structural and Normative Conditions for Interethnic Friendships in Multiethnic Classrooms. *International Journal of Intercultural Relations* 47. 1–12.

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

Snijders, T. A. B., Van de Bunt, G. G., and Steglich, C. E. G. (2010). Introduction to stochastic actor-based models for network dynamics. Social networks, 32(1):44-60.

Snijders, T. A. B. and Bosker, R. J. (2012). Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling. Sage, London, 2nd edition

Snijders, T. A. B. and C. Steglich. (2015). Representing Micro–Macro Linkages by Actor-based Dynamic Network Models. *Sociological Methods & Research* (44). 222–71.

Stadtfeld, C., Snijders, T.A.B., Steglich, C., and M.A.J. van Duijn. (2018a). Statistical Power in Longitudinal Network Studies. *Sociological Methods & Research.* Forthcoming.

Stadtfeld, C., Takács, K., and V., András. (2018b) The Emergence and Stability of Groups in Social Networks. Available at SSRN:

http://dx.doi.org/10.2139/ssrn.3232958 (accessed October 16, 2018)

Ripley, R.M., Snijders, T.A.B., Boda, Z., Vörös, A., and P. Preciado (2019). Manual for SIENA version 4.0. Oxford: University of Oxford, Department of Statistics; Nuffield College.

Rusbult, C. E. and Buunk, B. P. (1993). Commitment processes in close relationships: An interdependence analysis. *Journal of Social and Personal Relationships.* 10(2). 175-204.

Tajfel, H. and Turner, J. C. (1979). An integrative theory of intergroup conflict. *The social psychology of intergroup relations.* 33-47.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. Journal of Statistical Software, 36(3):1–48.

Wimmer, .; and Lewis, K. (2010): Beyond and Below Racial Homophily: ERG Models of a Friendship Network Documented on Facebook. *American Journal of Sociology* 116 (2), 583–642.

5 EPILOGE

5.1 Contribution and implications

This thesis focussed on the interrelation of ethnic homophily and endogenous network mechanisms as processes leading to ethnic segregation of classroom friendship networks. The theoretical framework clearly differentiates segregation as a characteristic of a network state which can result from homophily as a micro-level process. This clarification is necessary when investigating inter-dependent data such as friendship networks, because endogenous network mechanisms can lead to the emergence and stability of highly segregated networks even in the absence of strong homophily. In order to investigate homophily and it's interrelation with endogenous network mechanisms one has to look at network dynamics, rather than the existing network structures at a given time. This is because different network processes can lead to the same network structures. Therefore, if one wants to better understand how friendship networks in classrooms become ethnically segregated in order to identify strategies to counter-act this tendency, one has to take into account a variety of mechanisms that lead to how pupils are embedded into their surrounding network structures.

By applying the new hierarchical extension of stochastic actor oriented models, this study shows that transitivity – the tendency for friends-of-friends to be friends – differs within and between the group of native and non-native pupils. The analysis of 33 classrooms from Germany reveals that general homophily differs between groups, such that non-natives show a tendency for the formation and preservation of in-group friendships while natives do not. While both groups show a tendency for transitivity, natives' transitively embedded ties have a small tendency to rather be formed and maintained to the non-native out-group, who in turn show no such difference in

transitivity.

These findings seem initially counter-intuitive, given that cross-sectional descriptives show intragroup transitive triads to be the most prominent in our data. However, it is important to keep in mind that this cross-sectional observation does not allow us to draw conclusions about the dynamic process underlying the evolution of the networks. Relating the existing structure to natives' out-group tendency in transitivity suggests a possible explanation. During the early stages of the emerging network, or even before the pupils are assembled into the classroom, native pupils mainly have other native friends, leading to substantial in-group clustering. However, those native pupils who are friends with non-natives serve as facilitators, or bridges to the out-group. Native pupils' tendency to be friends with the friends of their friends overcomes potential group boundaries, leading to friendships with non-natives that are embedded in transitive triads. For the group of non-natives however, a different situation appears. Investigating the density of in- and out-group friendships by group, shows that during the first wave of data collection non-natives have nominated proportionally more natives than other non-natives as their friends. This was not the case for the group of natives, who show a stable tendency to proportionally nominate co-ethnics instead of non-natives. A possible explanation is that non-natives who named natives as their friends have not been nominated back, indicating an unbalanced relationship. These one-sided friendships are an unstable social relationships, due to the higher associated costs of between-group relations, especially when not counteracted by being embedded through a mutual friend. It can even be considered that natives reject the initial friendship nominations of natives. Following, it is not surprising to find that non-natives at the subsequent observation show to have a higher proportion of in- than out-group friendships. This suggests that the finding of non-natives driving the homophily effect, do so because ceasing to

consider natives they had previously nominated as friends and instead befriend other non-natives. These findings imply that non-natives initially are motivated to integrate themselves but stop their efforts when those are not acknowledged. Following, they might even build up an identification as non-natives.

Following this line of thought, the second study applies HSAOM to undertake a random effect analysis about the role of reciprocity for the creation and maintenance of pupils in- and out-group friendships in 194 Swedish classrooms. Despite the fact that in this work, homophily is not differentiated by groups, its findings support the implications of the previous work. Taken together, native and non-native pupils show homophily of friendship creation irrespective of these friendships being reciprocated or not. This supports the idea of pupils having to overcome group boundaries that increase subjectively associated costs of friendships with the out-group. Even if out-group members reciprocate friendship nominations, these boundaries are not overcome. For friendship maintenance however, reciprocity plays an important role especially between groups. Asymmetric friendships much rather sustain within groups, indicating that here the shared group identification can compensate pupils for their one-sided efforts of investing into the relationship with a co-ethnic, while group-differences make this scenario even less likely. Surprisingly, reciprocated friendships appear to be more stable between groups than within. Revealing heterophily of reciprocated friendship maintenance underlines that pupils perceive their relations to out-group members substantially different than their relations to the own group. A possible explanation is that balanced friendships between pupils from a different migration background demand a higher investment of both pupils in order to become established than between pupils of the same background. Thus, adolescents might perceive their reciprocated out-group friendships as more special by having overcome group boundaries. Friendships within

groups are not undergoing similar strain as regularly. Hence, they might not be considered as extraordinary, thus being more easily affected by negative outcomes of interactions such as disputes and will be dissolved more often.

While the previous studies focus on homophily two different endogenous network processes for the specific cases of either group differences or creation and maintenance of friendships, the goal of the last study is to combine tendencies of homophily for reciprocity, transitivity and centrality. Moreover, by applying data-driven agent based simulation, this work directly links the different micro-level mechanisms to the macro-outcome of segregation. Running single network SAOMs for 64 classrooms of the Swedish CILS4EU data without differentiating homophily by group or creation and maintenance, does not reveal any significant tendencies for homophily in endogenous network processes. However, the initially significant tendency for basic dyadic homophily becomes insignificant when accounting for in-group transitivity. Comparing the segregation levels of the networks simulated during the estimation to the observed values, does not support that accounting for homophily in endogenous network processes generally leads to better capturing the in- and out-group friendships in the network. Subsequently, the parameter estimates are used in agent-based simulations in order to test if the different homophily mechanisms are able to explain the emergence of the observed segregation starting from an empty network. The idea is that the process that led to the segregation at the first observation is modelled, under the assumption of estimated dynamics in network mechanisms being constant throughout the emergence of the first network. While the average segregation of the simulated networks performs slightly better when accounting for homophily in reciprocity, transitivity and centrality, these differences are not significantly different over all classrooms. However, since the simulations, on average produce higher than observed levels of segregation it follows

that accounting for homophily in reciprocity, transitivity and centrality represent desegregating mechanisms during the emergence of classroom friendship networks. While not explicitly modelled in this work, findings from the previous studies support this by finding that transitivity (at least for natives) contributes to creation and maintenance of friendship nominations between groups, and reciprocity being especially important for the maintenance of intergroup friendships.

In addition to the contributions made by investigating homophily in endogenous network effects and their link to segregation by cutting-edge statistical modelling techniques, the work on this thesis helped to further the development of technical aspects of conducting such research. Next to implementing several new effects that are now available for SAOM, a set of R-functions has been produced that helps future research on this topic by handling datasets with several networks (Appendix C) and produce dyad and triad census statistics for in- and out-group ties (Appendix D).

5.2 Discussion and conclusions

The substantive findings of this work provide insights into the mechanisms that drive pupils' friendship choices. While all presented studies show initial tendencies of pupils to rather befriend others with the same migration background, they underline that this is not to necessarily related to an actual in-group preference. The basic theoretical framework that is applied in this work did not have the goal to directly test theories about intergroup relations such as *group identity* or *contact theory*. However, it is possible to further extent the work presented here by integrating arguments from these theoretical approaches and derive testable hypotheses for different network mechanisms. Nonetheless, one goal of the conducted research was to better understand how the social social structures that pupils are embedded in, can influence their inter- and intragroup friendship choices. This was achieved by applying a rather general

theoretical framework based on *exchange theory* about a process in which friendships emerge from repeated interactions. In this context, baseline homophily is understood as an imbalance in pupils' perception of necessary investment, effort or risk in order to establish a social relation with classmates from their own or a different migration background.

One could argue that this approach relies on pupils to be aware of all previous investments and their respective gain from these and pupils being callously calculating when considering who they want to become friends with. However, it is not necessary for this work to understand pupils as such rational actors, but rather consider that their encounters with others leave them with an impression of more or less sympathy towards the other. With enough subsequent positively perceived encounters, they become increasingly certain about enjoying these interactions, making them more inclined to continue these. Or in the opposite case, negative encounters of simply unimpressive, awkward or even unpleasant exchanges with another, make pupils feel less inclined to continue spending time with them.

The decisions of each pupil about who they nominate as their friend can also be influenced by other pupils' friendship choices. Given the understanding of the friendship formation process, this does not rely on pupils explicitly knowing about others' friendship nominations, since the nominations are an outcome of repeated interactions with each other in which limited resources are exchanged. In the context of the classroom, pupils spend a considerable amount of their time with each other and thus many interactions will take place there making them likely to be observed by those not directly involved.

Based on the findings of the presented studies, different network mechanisms have been outlined that show an integrating effect for friendships between native and non-native

pupils. First of all, transitivity leads to group boundaries being overcome to the extent that pupils tend to become friends with their classmates irrespectively if they are from their in- or out-group so far as they share a mutual friend. Thus, if wanting to counter-act classrooms to become segregated, interventions should focus on establishing intergroup friendships as early as possible from the time that pupils are assembled into the classrooms. Potential interventions should be targeted to provide a setting in which pupils feel similarly comfortable interacting with co-ethnic peers or those of other background. Thus, if early on more relations are formed between groups, pupils also have a higher propensity to be indirectly connected to out-group alters For example, this could be achieved by trying to counter-act existing negative out-group attitudes or prejudice, or encouraging and implementing activities where pupils with different backgrounds interact, cooperate and develop a basis to identify with each other. If in the early emergence of the friendship network indirect links between pupils of different background develop, tendencies for clustering would not be based on ethnicity. Based on the finding that mutual friendships tend to be maintained especially between groups, an early encouragement of intergroup friendships to become reciprocated can help to prevent networks from segregating at later stages. However, the scope of this work does not allow to derive approaches for designing interventions.

The research presented here is meant to provide a foundation for theoretical and empirical work that aims to understand causes of ethnic segregation in friendship networks. It showed that accounting for different network structures and their tendencies of change, one can gain valuable insights about how the positioning of pupils within these structures is related to their choices of whom to be friends with. This approach can be utilised and expanded by future work, in order to refine existing theories that explain homophily and the emergence of segregations in order to

incorporate mechanisms that are inherent to the dynamics social networks. One way to build on this work could be to incorporate tendencies for homophily in the three endogenous network mechanisms simultaneously, while differentiating them for creation and maintenance of friendships. This would allow us to more clearly understand how each of these mechanisms contributes to either the emergence or dynamic of segregation. Another way to improve on the presented work is to relax some of the assumptions made due to modelling choices. One could be to directly for the role of social foci by collecting data beyond the classroom and school about pupils shared leisure time activities and club memberships. Accounting for shared membership in such ventures when analysing ethnic homophily should allow to more clearly map the forces that drives pupils in- and out-group friendship choices. A further valuable approach would be to investigate time heterogeneity mechanisms that drive friendship networks. This work assumes that during the initial stages of emerging networks, mechanisms of friendship creation are the major driving force of network evolution. Gaining more insight about how different network mechanisms might influence pupils friendship creation early on, can provide valuable grounds for further theory development and policy suggestions. Additionally the simplification of pupils being either from native or non-native backgrounds. Accounting for different ethnic origins would allow to more clearly account for arguments about ethnic identification as possible differences between groups. Moreover, when having information on multiple classrooms, applying HSAOMs can test if the heterogeneity of homophily between networks is systematically related to the the network context, such as the classroom composition or even initial ethnic diversity in classrooms. Ideally, the research direction of this study is the beginning of more intergroup research that explicitly takes into account the nature of social relations to be embedded in social structures.

6 APPENDIX A

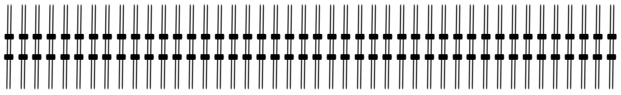


Table A1: HSAOM full model specification

7 APPENDIX B

Effect names	M1		M2	
	b	s.e.	b	s.e.
Core effects (maintenance)				
reciprocity	3.339	.433 ***	4.265	.572 ***
egoX_migrant	-.328	.277	-.062	.249
altX_migrant	.328	.170 *	.065	.190
sameX_migrant	.241	.133 *	.599	.217 ***
sameX_migrant X reciprocity			-.844	.230 ***
Core effects (creation)				
reciprocity	1.022	.268 ***	.697	.559
egoX_migrant	0.05	.505	-.472	.404
altX_migrant	-.385	.241 *	-.048	.258
sameX_migrant	.343	.192 *	.017	.388
sameX_migrant X reciprocity			.952	.351 ***
Homophily controls (evaluation)				
egoX_female	.050	.118	.065	.112
altX_female	.018	.106	-.001	.114
sameX_female	.453	.103 ***	.495	.114 ***
egoX_smoke	-.192	.207	-.501	.416
altX_smoke	-.323	.158 **	-.421	.191 *
sameX_smoke	.172	.145	.080	.140
egoX_traditional family	.251	.152 *	.276	.177 *
altX_traditional family	.008	.082	-.023	.101
sameX_traditional family	.059	.071	.037	.072
egoX_ISEI	.002	.032	.003	.034
altX_ISEI	.001	.032	.002	.033
ISEI similarity	.241	.135 *	.277	.136 *
egoX_age	-.061	.043	-.075	.046
altX_age	-.016	.038	-.028	.039
age similarity	.086	.091	.097	.066
egoX_language ability	.034	.038	.038	.041
altX_language ability	.014	.035	.005	.035
language ability similarity	-.495	.216 *	-.343	.160 **
egoX_cognitiy ability	.003	.008	.003	.007
altX_cognitive ability	-.018	.005 ***	-.015	.005 ***
cognitive ability similarity	.194	.150	.180	.160
egoX_friends out of class	-.122	.070 *	-.141	.078 *
Network controls				
transitive triplets	1.465	.180 ***	1.572	.254 ***
reciprocated transitive triplets	-.580	.117 ***	-.654	.126 ***
indegree popularity	-.218	.068 ***	-.258	.117 *
outdegree popularity	.040	.093	.151	.175
indegree activity	-.058	.135	-.217	.202
outdegree activity	-.441	.086 ***	-.412	.089 ***

Table B1: HSAOM full model specification

8 APPENDIX C

8.1 Function to generate a list of siena datasets from surveydata

SienaShape_v04.R

```
# v01: basic functionality
# v02: include behavior variable
# v03: include multiplex networks
# v04: include composition change object and options maxlike | bayes



classnets <-
  function(classes,
           idvar,
           netvar1,
           netvar2,
           nvar,
           multiplex1 = NULL,
           multiplex2 = NULL,
           mvar1 = NULL,
           mvar2 = NULL,
           depbeh = NULL,
           nwave = NULL,
           comp = "structural",
           covars,
           join = NULL,
           leave = NULL,
           jlval = 10,
           missings = NULL,
           missval = NULL) {
    library(RSiena)
    netshape <-
      function(classnr,
               egovar,
               netvar,
               nvar,
               structural = NULL,
               strval = NULL,
               miss = NULL,
               mval = NULL) {
        idlist <- na.omit(egovar[substr(egovar, 1, 6) == as.character(classnr)])
        adj <- matrix(0, nrow = length(idlist), ncol = length(idlist))
        netvars <- paste(netvar, c(1:nvar), sep = "_")
        strmiss <- structural[substr(egovar, 1, 6) == as.character(classnr)]
```

```
  for (v in netvars) {
   netnom <-
    eval(parse(text = v))[substr(egovar, 1, 6) == as.character(classnr)]
   nom <- match(netnom, idlist, nomatch = 0)
   ind <- c(1:length(idlist))

   for (i in 1:length(nom)) {
    if (nom[i] != 0) {
     adj[[ind[i], nom[i]]] <- 1
    }
   }
  }

  for (i in 1:length(idlist)) {
   if (is.null(strmiss) == FALSE) {
    if (strmiss[i] == 1 && is.na(strmiss) == FALSE) {
     adj[i, ] <- strval
     adj[, i] <- strval
    }
   }
  }

  return(adj)
 }

siena.cov <- function(classnr, idvar, covar) {
 a <- covar[substr(idvar, 1, 6) == as.character(classnr)]
 cocov <- coCovar(a, centered = FALSE)
 return(cocov)
}

behavar <- function(classnr, egovar, depbeh, nwave) {
 for (i in 1:nwave) {
  if (i == 1) {
   behvar <-
    as.matrix(eval(parse(text = paste(
     depbeh, i,
     sep = ""
    )))[substr(egovar, 1, 6) == as.character(classnr)])
  }
  else {
   behvar <-
    cbind(behvar, eval(parse(text = paste(
     depbeh, i,
     sep = ""
    )))[substr(egovar, 1, 6) == as.character(classnr)])
  }
 }

 behavior <- sienaDependent(behvar, type = "behavior")
```

```
}

nets1 <- list()
nets2 <- list()
dimension <- list()
cocov <- list()
sienadep <- list()
network <- list()

if (is.null(multiplex1) == FALSE && is.null(multiplex2) == FALSE) {
 sienadep2 <- list()
}

for (j in 1:length(classes)) {
 nets1 <-
  netshape(
   classes[j],
   idvar,
   netvar1,
   nvar,
   structural = join,
   strval = jlval,
   miss = missings,
   mval = missval
  )
 nets2 <-
  netshape(
   classes[j],
   idvar,
   netvar2,
   nvar,
   structural = leave,
   strval = jlval,
   miss = missings,
   mval = missval
  )

 # for maximum likelihood and bayes, exclude inactive cases at first observation
 if (comp == "maxlike" | comp == "bayes") {
  nets1[which(nets1 == 10)] <- NA
 }

 dimension <- c(dim(nets1), 2)
 sienadep <- sienaNet(array(c(nets1, nets2), dim = dimension))

 if (is.null(multiplex1) == FALSE && is.null(multiplex2) == FALSE) {
  nets1 <-
   netshape(
    classes[j],
    idvar,
```

```
      multiplex1,
      mvar1,
      structural = join,
      strval = jlval,
      miss = missings,
      mval = missval
    )
  nets2 <-
    netshape(
      classes[j],
      idvar,
      multiplex2,
      mvar2,
      structural = leave,
      strval = jlval,
      miss = missings,
      mval = missval
    )

  dimension <- c(dim(nets1), 2)
  sienadep2 <- sienaNet(array(c(nets1, nets2), dim = dimension))
}

covs <- list()
input <- ""
for (cov in 1:length(covars)) {
  if (cov == 1) {
    input[[cov]] <- paste(names(covars[cov]), "=", "covs[[", cov, "]]",
      sep =
        ""
    )
  }
  else {
    input <-
      c(
        input,
        paste(",", names(covars[cov]), "=", "covs[[", cov, "]]", sep = "")
      )
  }

  covs[[cov]] <- siena.cov(classes[j], idvar, covars[[cov]])
}

if (is.null(depbeh) == FALSE && is.null(nwave) == FALSE) {
  behave <- behavar(classes[[j]], idvar, depbeh, nwave)

  network[[j]] <-
    eval(parse(text = (
      paste(
        "sienaDataCreate(netw=sienadep,",
```

```
            depbeh,
            "=behave,",
            paste(input, collapse = ""),
            ")"
          )
        )))
      }
      if (is.null(multiplex1) == FALSE && is.null(multiplex2) == FALSE) {
        network[[j]] <-
          eval(parse(text = (
            paste(
              "sienaDataCreate(netw1=sienadep,netw2=sienadep2,",
              paste(input, collapse = ""),
              ")"
            )
          )))
      }
      else {
        network[[j]] <-
          eval(parse(text = (
            paste(
              "sienaDataCreate(netw=sienadep,",
              paste(input, collapse = ""),
              ")"
            )
          )))
      }
    }

    return(network)
  }
```

8.2 Functions to select networks suitable for single classroom SAOMs

SienaSelect_v03.R

```
# V.1: basics
    # v.2: solving issues with structural missings
    # v.3: Fixing Differences to print01report from treating structural missings as 0 in
    both waves


    SienaSelect <- function(netlist) {
     # Identify structural missings to exclude them for jaccard calculation
     strmiss <- function(var) {
     varmiss <- which(rowSums(is.na(var)) == ncol(var))
     return(varmiss)
     }
     covprop <- function(net) {
     covarprop <- matrix(rep(NA, length(net$cCovars)), ncol = length(net$cCovars))
```

```
colnames(covarprop) <- names(net$cCovars)
for (n in 1:length(net$cCovars)) {
covarprop[, n] <- attributes(net$cCovars[[n]])$mean > 0 &&
attributes(net$cCovars[[n]])$mean < 1
}
return(covarprop)
}
netstats <- function(net) {
# First create empty results objects
avdeg <- matrix(rep(NA, net$observation * length(net$dep)), nrow =
length(net$dep))
up <- matrix(rep(NA, (net$observations - 1) * length(net$dep)), nrow =
length(net$dep))
down <- matrix(rep(NA, (net$observations - 1) * length(net$dep)), nrow =
length(net$dep))
# For jaccard calculation, number of periods necessary
periods <- net$observations - 1
jaccard_exp <- matrix(rep(NA, periods), nrow = length(net$dep))
jaccard_obs <- matrix(rep(NA, periods), nrow = length(net$dep))

# Run selection diagnostics for each dependent networkvariable
for (i in 1:length(net$dep)) {
# Average degree
for (j in 1:net$observations) {
avdeg[i, j] <- mean(attributes(net$dep[[i]])$degree[j])
}

# Jaccard
for (p in 1:periods) {
var1 <- net$depvars[[i]][, , p]
var2 <- net$depvars[[i]][, , p + 1]

    # set joiners/leavers to 0 in both waves
    var1[which(var1 == 10)] <- 0
    var1[which(var2 == 10)] <- 0

    var2[which(var2 == 10)] <- 0
    var2[which(var1 == 10)] <- 0

    for (k in 1:nrow(var1)) {
      for (l in 1:nrow(var1)) {
        if (is.na(var1[l, k]) == TRUE && is.na(var2[l, k]) == TRUE) {
          var2[l, k] <- var1[l, k]
        }
      }
    }
    if (sum(dim(table(var1, var2))) == 4) {
      tab <- table(var1, var2)
      pa <- tab["1", "1"] + tab["1", "0"]
      pb <- tab["1", "1"] + tab["0", "1"]
```

```
        pab_ind <- pa * pb / sum(tab)
        jaccard_exp[i, p] <- pab_ind / (pa + pb - pab_ind)
        jaccard_obs[i, p] <- tab["1", "1"] / (tab["1", "1"] + tab["0", "1"] + tab["1",
"0"])
       }
       else {
        jaccard_exp[i, p] <- NA
        jaccard_obs[i, p] <- NA
       }
      }

     # Up or down only
     up[i, ] <- attributes(net$dep[[i]])$uponly
     down[i, ] <- attributes(net$dep[[i]])$downonly
    }

    net.select <- list(average.degree = avdeg, expected.jaccard = jaccard_exp,
observed.jaccard = jaccard_obs, up.only = up, down.only = down)

    return(net.select)
  }


  sel.stats <- list()
  for (n in 1:length(netlist)) {
    net <- netlist[[n]]
    sel.stats[[n]] <- netstats(net)
    sel.stats[[n]]$cCovprop <- covprop(net)
  }
  return(sel.stats)
}

which.keep <- function(keep.from, networks, degree.min = 1.5, degree.max = 15,
jacc.min = 0.3, jacc.max = 0.98) {
  keeplist <- rep(NA, length(keep.from))

  for (i in 1:length(networks)) {
    keep <- NA
    # first check for migrant composition
    if (sum(keep.from[[i]]$cCovprop[1:2]) != 2 |
      (keep.from[[i]]$down + keep.from[[i]]$up != 0) |
      (is.na(sum(keep.from[[i]]$average.degree)) == TRUE) |
      (is.na(keep.from[[i]]$observed.jaccard) == TRUE) |
      (min(keep.from[[i]]$average.degree) < degree.min) |
      (max(keep.from[[i]]$average.degree) > degree.max) |
      (keep.from[[i]]$observed.jaccard < jacc.min &&
        keep.from[[i]]$observed.jaccard >= jacc.max) == TRUE) {
      keep <- FALSE
    } else {
      keep <- TRUE
```

```
    }
    keeplist[i] <- keep
  }

  return(keeplist)
}
```

8.3 Function to manipulate existing siena objects

sienaRecode_v102.R

```
#v01 - basic functionality with constant covariates
    #v02 - all covariate types
    #v03 - add option 'addCovs' to extract covariates from one siena object and create
    sienaData with new network
    #v04 - Restructure function in order to work with existing sienaRecode instead of
    changing sienaDataCreate function
    #v1.0 - changed option addcov to option action with the choices of:
    recode,addcov,newdep and addep
    #     - Replace dependent and add dependent
    #v1.01- supports multiple dependent variables
    #v1.02- support multigroup objects

    sienaRecode <-
      function(sdata,
               newvar,
               operation = NULL,
               options = NULL,
               constant = TRUE,
               dyadic = FALSE,
               action = 'recode',
               newob = NULL,
               ...) {
      #dots<-list(...)
      if (action != 'recode' &&
          action != 'addcov' && action != 'newdep' && action != 'addep') {
        cat('Need valid action from: recode, addcov, newdep or addep')
        break

      }
      #Preserve original input
      input <- operation

      if (is.null(options) == F) {
        options <- paste(',', options, collapse = "")
      }

      addCovs <- function(sdata) {
        covariates <- list()
```

```
  cov.names <- list()
  covstring <- NULL
  for (i in c("cCovars", "vCovars", "dycCovars", "dyvCovars")) {
    cov.names[[i]] <- names(sdata[[i]])
    if (length(cov.names[[i]]) > 0) {
      covstring[[i]] <-
        paste("",
              cov.names[[i]],
              "'=sdata[['",
              i,
              "']]$",
              cov.names[[i]],
              ",",
              collapse = "")
    }
  }
  covstring <- paste(covstring, collapse = '')
  covstring <- substring(covstring, 1, nchar(covstring) - 1)
  return(covstring)
}
addDeps <- function(sdata) {
  #Extract original sienaData dependent
  depstring <- NULL
  for (i in names(sdata$depvars)) {
    depstring[[i]] <-
      paste(i, "= sdata$depvars[['", i, "']],", collapse = "")
  }
  depstring <- paste(depstring, collapse = '')

  return(depstring)
}

if (class(sdata)[1] == "sienaGroup") {
  runs <- length(sdata)
  groupdata <- sdata
  netlist <- list()
  group.ind <- T
}
else {
  runs <- 1
  group.ind <- F
}



for (group in c(1:runs)) {
  if (group.ind == T) {
    sdata <- groupdata[[group]]
  }
```

```
  if (action == 'recode') {
    if (is.null(operation) == T) {
      cat('Some sort of recoding operation needs to be specified')
      break
    }
    # First check type of variables relevant for recode and extract them
    # Which variables are used in operation
    covariates <- list()
    cov.names <- list()

    #Get covariates and names used in operation from existing siena object
    for (i in c("cCovars", "vCovars", "dycCovars", "dyvCovars")) {
      covariates[[i]] <-
        which(lapply(names(sdata[[i]]), grep, operation) >= 1)
      cov.names[[i]] <-
        names(sdata[[i]])[which(lapply(names(sdata[[i]]), grep, operation) >= 1)]

    }


    if (constant == T) {
      #Replace variable name in command line with reference to variable in siena
data object
      for (covs in c("cCovars", "vCovars", "dycCovars", "dyvCovars")) {
        if (length(covariates[[covs]]) > 0) {
          for (i in 1:length(covariates[[covs]])) {
            operation <-
              gsub(
                cov.names[[covs]][i],
                paste(
                  'sdata[[',
                  paste("'", covs, "'", sep = ""),
                  ']][[',
                  covariates[[covs]][i],
                  ']]',
                  collapse = ""
                ),
                operation
              )
          }
        }
      }

      #Execute recode command line
      eval(parse(text = paste(newvar, '<-', operation)))

      #And create new siena variable to be added to siena data
      if (dyadic == T) {
        eval(parse(text = paste(
          'attributes(', newvar, ')<-NULL', sep = ""
```

```
  )))
  if (group.ind == T) {
    shape <- c(length(sdata$nodeSets$Actors),
               sdata$observations - 1)
    eval(parse(
      text = paste(
        newvar,
        '<-matrix(',
        newvar,
        ",",
        paste(shape, collapse = ","),
        ')',
        sep = ""
      )
    ))
    eval(parse(
      text = paste(
        newvar,
        '<-varDyadCovar(',
        newvar,
        options,
        ')',
        sep = ""
      )
    ))
  }
  else{
    eval(parse(
      text = paste(
        newvar,
        '<-coDyadCovar(matrix(',
        newvar,
        ',nrow=length(sdata$nodeSets$Actors))',
        options,
        ')',
        sep = ""
      )
    ))
  }
}
if (dyadic == F) {
  eval(parse(text = paste(
    'attributes(', newvar, ')<-NULL', sep = ""
  )))
  if (group.ind == T) {
    shape <- c(length(sdata$nodeSets$Actors),
               sdata$observations - 1)
    eval(parse(
      text = paste(
        newvar,
```

```
        '<-matrix(',
        newvar,
        ",",
        paste(shape, collapse = ","),
        ')',
        sep = ""
      )
    ))
    eval(parse(
      text = paste(newvar, '<-varCovar(', newvar, options, ')', sep = "")
    ))
  }
  else{
    eval(parse(
      text = paste(newvar, '<-coCovar(', newvar, options, ')', sep = "")
    ))
  }
 }


}


if (constant == F) {
  #Preserve original input line
  input <- operation
  observations <- sdata$observations

  #First replace constant covariates in command line
  for (covs in c("cCovars")) {
    if (length(covariates[[covs]]) > 0) {
      for (i in 1:length(covariates[[covs]])) {
        operation <-
          gsub(
            cov.names[[covs]][i],
            paste(
              'sdata[[',
              paste("'", covs, "'", sep = ""),
              ']][[',
              covariates[[covs]][i],
              ']]',
              collapse = ""
            ),
            operation
          )

      }
    }
  }
```

```
        #Since new outcome is varying, final command line needs to be executed for
each measurement
        operation <- list(operation, operation)
        if (dyadic == F) {
          for (covs in c("vCovars", "dyvCovars")) {
            if (length(covariates[[covs]]) > 0) {
              for (i in 1:length(covariates[[covs]])) {
                for (period in 1:(sdata$observations - 1)) {
                  operation[[period]][1] <-
                    gsub(
                      cov.names[[covs]][i],
                      paste(
                        'sdata[[',
                        paste("'", covs, "'", sep = ""),
                        ']][[',
                        covariates[[covs]][i],
                        ']][,',
                        period,
                        ']',
                        collapse = ""
                      ),
                      operation[[period]][1]
                    )
                }
              }
            }
          }
          newvarname <- newvar
          #Create empty object in order to execute command:
          eval(parse(
            text = paste(
              newvar,
              '<- matrix(rep(NA,length(sdata$nodeSets$Actors)*(observations-
1)),ncol=(observations-1))',
              sep = ""
            )
          ))


          for (period in 1:(sdata$observations - 1)) {
            eval(parse(text = paste(
              newvarname, '[,', period, ']<-', operation[[period]]
            )))
          }

          eval(parse(
            text = paste(newvar, '<-varCovar(', newvar, options, ')', sep = "")
          ))
        }
```

```
        if (dyadic == T) {
          for (covs in c("dyvCovars")) {
            if (length(covariates[[covs]]) > 0) {
              for (i in 1:length(covariates[[covs]])) {
                for (period in 1:(sdata$observations - 1)) {
                  operation[[period]][1] <-
                    gsub(
                      cov.names[[covs]][i],
                      paste(
                        'sdata[[',
                        paste("""", covs, """, sep = ""),
                        ']][[',
                        covariates[[covs]][i],
                        ']][,,',
                        period,
                        ']',
                        collapse = ""
                      ),
                      operation[[period]][1]
                    )
                }
              }
            }
          }
          newvarname <- newvar
          #Create empty object in order to execute command:
          #     vdyc<-
varDyadCovar(array(c(t(friendship[,,1]),t(friendship[,,2]),2),dim=c(dim(t(friendsh
ip[,,1])),2)))

          eval(parse(
            text = paste(
              newvar,
              '<-
array(c(rep(matrix(rep(NA,length(sdata$nodeSets$Actors)^2),ncol=(observations-
1)),2),observations-
1),dim=c(length(sdata$nodeSets$Actors),length(sdata$nodeSets$Actors),2))',
              sep = ""
            )
          ))


          for (period in 1:(sdata$observations - 1)) {
            eval(parse(
              text = paste(newvarname, '[,,', period, ']<-', operation[[period]])
            ))
          }

          eval(parse(
            text = paste(newvar, '<-varDyadCovar(', newvar, options, ')', sep = "")
```

```
    ))
  }

}

#Now create string to add original covariates
covstring <- addCovs(sdata)
#And dependent
depstring <- addDeps(sdata)
#Generate commandline to create new siena dataset
datastring <-
  paste(
    'z <- sienaDataCreate(',
    depstring,
    paste("", newvar, "="),
    newvar,
    ",",
    covstring,
    ')',
    sep = ""
  )

}
if (action == 'addcov') {
  if (class(newob) != 'coCovar' && class(newob) != 'varCovar') {

  }
  #Now create string to add original covariates
  covstring <- addCovs(sdata)
  #and dependent variables
  depstring <- addDeps(sdata)
  #Newly create siena data for that variable
  datastring <-
    paste(
      'z <- sienaDataCreate(',
      depstring,
      paste("", newvar, "=newob,"),
      covstring,
      ')',
      sep = ""
    )
}
if (action == 'newdep') {
  if (is.null(newob) == T) {
    cat('Need new siena Dependent variable')
    break
  }

  if (class(newob) != 'sienaDependent') {
    cat('New dependent variable not sienaDependent format')
```

```
    break
   }
   covstring <- addCovs(sdata)
   depstring <- paste("'", newvar, "'= newob,")
   datastring <-
    paste('z <- sienaDataCreate(', depstring, covstring, ')', sep = "")
  }
  if (action == 'addep') {
   if (is.null(newob) == T) {
    cat('Need new siena Dependent variable')
    break
   }

   if (class(newob) != 'sienaDependent') {
    cat('New dependent variable not sienaDependent format')
    break
   }
   covstring <- addCovs(sdata)
   depstring <- addDeps(sdata)

   datastring <-
    paste(
     'z <- sienaDataCreate(',
     depstring,
     paste("'", newvar, "'= newob,"),
     covstring,
     ')',
     sep = ""
    )



  }


  #Execute respectively created command line to create new siena dataset
  eval(parse(text = gsub(" ", "", datastring)))

  if (group.ind == T) {
   netlist[[group]] <- z
  }

 }

 #Final report
 if (action == 'newdep') {
  print(paste(
   "New Siena Data created with:",
   newvar,
   " as new dependent variable",
```

```
    sep = " "
  ))
 }
 if (action == 'addep') {
  print(
   paste(
    "New Siena Data created with:",
    newvar,
    " as dependent additional dependent variable",
    sep = " "
   )
  )
 }
 if (action == 'recode') {
  print(paste("New variable:", newvar, " is created as:", input, sep = " "))
 }
 if (action == 'addcov') {
  print(paste("New variable:", newvar, " is added to the dataset", sep = " "))
 }
 if (group.ind == T) {
  z <- sienaGroupCreate(netlist)
 }
 return(z)

}
```

9 Appendix D

9.1 Function to change dataformat

netprep.R

```
netprep <- function(netobject, netvar, dicocov) {
      networks <- list()
      if (class(netobject) == "siena") {
        # Basic data characteristics
        # number of ego's
        nego <- length(netobject$nodeSets$Actors)

        # proportion of covar=1
        propcov <- attributes(netobject$cCovars[[dicocov]])["mean"]

        ccov <- netobject$cCovars[[dicocov]]
        # Get covariate object
        cov.nomiss <- which(is.na(ccov) == F)
        ccov <- ccov[cov.nomiss]

        # Number of timepoints
        obs <- netobject$observations

        # Get network matrices
        for (i in 1:obs) {
          networks[[i]] <- netobject$depvars[[netvar]][, , i]

          # Set structural missings to NA
          networks[[i]][which(networks[[i]] == 10)] <- NA
          # Unly take network with covariate values
          networks[[i]] <- networks[[i]][cov.nomiss, cov.nomiss]
        }

        actor.list <- list()
        for (i in 1:obs) {
          actor.list[[i]] <- apply(networks[[i]], 1, function(x) length(which(is.na(x))) <
    length(x))
        }
        balanced.actors <- which(Reduce("+", actor.list) == obs)
        for (i in 1:obs) {
          networks[[i]] <- networks[[i]][balanced.actors, balanced.actors]
        }
        ccov <- ccov[balanced.actors]

        networks
      }

      return(list(networks = networks, covariate = ccov, index = balanced.actors))
```

```
}
```

9.2 Function to calculate triad cencus by group-composition

triadcount_v01.R

```
triad.groups <- function(net1, net2, att, wishlist1 = NULL, wishlist2 = NULL, na.val =
    9999.99, aggregate = "actor.total") {
     if (aggregate != "actor.total" && aggregate != "number.alters" && aggregate !=
    "network") {
       print("Wrong aggregation method specified")
       break
     }

     # colnames(triads)<-c('ih','hi','ij','ji','hj','jh')



     # Find all existing covariat categories
     cat.vals <- as.integer(attributes(table(att, exclude = NULL))$dimnames$att)
     cat.names <- paste(cat.vals)
     wishes.att <- as.matrix(expand.grid(
       c(cat.vals), c(cat.vals), c(cat.vals)
     ))

     att[which(is.na(att) == T)] <- na.val

     if (is.null(wishlist1) == T) {
       wishlist1 <- c(2, 2, 2, 2, 2, 2)
       # wishlist2 <- c(1,1,1,0,1,1)
     }
     if (is.null(wishlist2) == T) {
       wishlist2 <- c(2, 2, 2, 2, 2, 2)
     }

     # Create relational covariate matrix
     vec1 <- matrix(1, 1, length(att))
     mat.diag <- matrix(1, nrow(net1), nrow(net1))
     diag(mat.diag) <- 0
     mat.sa <- att %*% vec1

     # Create empty results object
     n.att <- length(cat.names)
     # For aggregate level: individual
     if (aggregate == "actor.total" | aggregate == "number.alters") {
       triad.census <- array(NA, c(length(att), n.att, n.att), list(paste(c(1:length(att)),
    sep = ""), cat.names, cat.names))
     }
     # For aggregate level: network
     if (aggregate == "network") {
       triad.census <- array(NA, c(n.att, n.att, n.att), list(cat.names, cat.names,
```

```
cat.names))
  }
  for (comp in c(1:nrow(wishes.att))) {
    wishlist.att <- wishes.att[comp, ]
    wishlist.att[which(is.na(wishlist.att) == T)] <- na.val

    tempmat <-
      (
        (
          ((
            (((net1 == wishlist1[1]) + (wishlist1[1] == 2)) * ((net2 == wishlist2[1]) +
(wishlist2[1] == 2))
              * ((t(net1) == wishlist1[2]) + (wishlist1[2] == 2)) * ((t(net2) ==
wishlist2[2]) + (wishlist2[2] == 2)))
            *
              ((t(mat.sa) == wishlist.att[2]))) * mat.diag)
          %*%
            ((((net1 == wishlist1[5]) + (wishlist1[5] == 2)) * ((net2 == wishlist2[5]) +
(wishlist2[5] == 2)) *
              ((t(net1) == wishlist1[6]) + (wishlist1[6] == 2)) * ((t(net2) ==
wishlist2[6]) + (wishlist2[6] == 2))) * mat.diag)
        ) *
          ((((net1 == wishlist1[3]) + (wishlist1[3] == 2)) * ((net2 == wishlist2[3]) +
(wishlist2[3] == 2)) *
            ((t(net1) == wishlist1[4]) + (wishlist1[4] == 2)) * ((t(net2) == wishlist2[4])
+ (wishlist2[4] == 2)))
          *
            ((mat.sa == wishlist.att[1]) * (t(mat.sa) == wishlist.att[3])) * mat.diag)
      )

    diag(tempmat) <- NA
    if (aggregate == "network") {
      triad.census[paste(wishes.att[comp, 1]), paste(wishes.att[comp, 3]),
paste(wishes.att[comp, 2])] <- sum(tempmat, na.rm = T)
    }
    if (aggregate == "number.alters") {
      triad.census[which(att == wishes.att[comp, 1]), paste(wishes.att[comp, 3]),
paste(wishes.att[comp, 2])] <- apply(tempmat, 2, function(x) {
        length(which(x != 0))
      })[which(att == wishes.att[comp, 1])]
    }
    if (aggregate == "actor.total") {
      triad.census[which(att == wishes.att[comp, 1]), paste(wishes.att[comp, 3]),
paste(wishes.att[comp, 2])] <- apply(tempmat, 2, sum, na.rm = T)[which(att ==
wishes.att[comp, 1])]
    }
  }

  attributes(triad.census)["wishlist1"] <- as.character(paste(wishlist1, collapse =
","))
```

```
  attributes(triad.census)["wishlist2"] <- as.character(paste(wishlist2, collapse =
","))
  class(triad.census) <- "triad.census"
  return(triad.census)
}
```

10 Appendix E – Statements of co-authorship

CO-AUTHORSHIP DECLARATION FORM

1. Student to complete:

Name: Robert Hellpap

Thesis Title: The role of endogenous network processes for ethnic friendship segregation in classrooms

Title of Chapter which is co-authored: Choice and structural homophily for native and non-native pupils in classroom friendship network

2. Co-Author to complete:

Name: Dr. Enis Bicer, Akkon-Hochschule für Humanwissenschaften, Berlin, Germany

Describe your contribution to this chapter:

During the course of working on the paper, Enis Bicer's responsibility was mainly the theoretical frame and explication of hypotheses. He also is co-responsible for the introduction and conclusion of the paper.

"I hereby declare that this chapter is predominantly original work by the student named above"

Signature & Date: 4.7v.2018

3. Supervisor (if different from co-author)

Name:

Signature & Date:

Date approved:

Signature of DGS:

E.1 - Co-authorship declaration: Dr. Enis Bicer

Co-Authorship Declaration Form

1. Student to complete:

Name: Robert Hellpap

Thesis Title: Investigating the role of network inherent processes for ethnic friendship segregation in classrooms.

Title of Chapter which is co-authored: Who becomes and who stays friends? Ethnic segregation as a result of group-specific creation and dissolution of friendship in classrooms

2. Co-Author to complete:

Name: Isabel Raabe

Describe your contribution to this chapter:

Isabel contributes mainly to the literature review, outlining of the general theoretical framework as well as with refining the implications of the results.

"I hereby declare that this chapter is predominantly original work by the student named above"

Signature & Date: 3/10/18 I. Raabe

3. Supervisor (if different from co-author)

Name:

Signature & Date:

Date approved:

Signature of DGS:

E.2 - Co-authorship declaration: Dr. Isabell Raabe

Co-Authorship Declaration Form

1. Student to complete:

Name: Robert Hellpap

Thesis Title: Investigating the role of network inherent processes for ethnic friendship segregation in classrooms.

Title of Chapter which is co-authored: Who becomes and who stays friends? Ethnic segregation as a result of group-specific creation and dissolution of friendship in classrooms

2. Co-Author to complete:

Name: Prof. Jan O. Jonsson

Describe your contribution to this chapter:

Prof. Jan O. Jonsson contributes mainly to the overall direction of the paper with regard to publication strategy, refining of theoretical arguments and implications of finding as well as language use.

"I hereby declare that this chapter is predominantly original work by the student named above"

Signature & Date: 2018, Sept 24th

3. Supervisor (if different from co-author)

Name:

Signature & Date:

Date approved:

Signature of DGS:

E.3 - Co-authorship declaration: Prof. Jan O. Jonsson