Social Aspects of Educational Inequality

Isabel Jasmin Raabe
Nuffield College
Department of Sociology
University of Oxford

A thesis presented for the degree of
Doctor of Philosophy
Hilary Term 2018

To my family

Acknowledgements

First of all, I want to thank my supervisor Jan O. Jonsson, who did not only manage to keep the fine balance between leaving me the freedom to pursue my interests and keeping me on track with my PhD thesis, but who also taught me that it is a lot easier to be a researcher when you keep your sense of humour. Thank you for your close supervision, giving detailed feedback and in-depth discussions, for collaborating, and for helping out with the financial means to attend conferences or to visit Nuffield from Zürich.

Furthermore, I would like to thank Christoph Stadtfeld and the other members of the Chair of Social Networks at the ETH Zürich for inviting me to the group, and for teaching me everything from how to use *R* to the application of state-of-the-art network methods, through group discussions, whiteboard sessions, and hand-drawn histograms.

I owe thanks to those who collaborated with me on chapters of this thesis, Ralf Wölfer, Christoph Stadtfeld, and Zsófia Boda, as well as to Robert Hellpap, who offered a lot of practical help at the beginning of my doctoral studies.

The financial resources from the Economic and Social Research Council and Nuffield College enabled me to engage in doctoral studies in the first place, and made it possible for me to attend winter schools, workshops, and conferences. This thesis benefitted vastly from these opportunities, and I want to thank the teachers of various methods courses and those who provided valuable feedback to my work.

There are several people who equipped me with the necessary skills and mind-set to pursue a PhD, and to whom I am truly grateful. Michael Biggs, my thesis supervisor in the M.Sc. in Oxford, who taught me the first things about how to write a sociological paper, who helped me digest the course material through further discussion, and who offered advice and practical help with PhD applications. Silke Stroh and Paul Spickard, for making me excited about academic work during the M.A. in Münster, for their advice, and for writing numerous reference letters. My old English teacher, Uwe Carstens, who did not only cast the seeds of my English language skills but who became my fatherly friend in the last decade after graduation from high school. Thank you for your encouragement, your advice, and your help with proofreading this thesis.

Apart from the academic support, I received a lot of backing and encouragement from my friends and my family, who made the process of writing a PhD thesis lighter and more enjoyable, even when they where not geographically close.

Gesa, thank you for your loyal friendship, for your frequent visits, and for finding ways to include me in your daily life even though I live in a different country. Jörn, for being the caring, funny, and wise friend that you are. Claudia, for your attentive friendship, your sincerity, and all your visits in the last five years. Andy, Lukas, David, Kati, Bine, and Alex – thank you for making every time we see each other feel like no time has passed since high school.

I am truly grateful for the friends I found in Oxford. Rosie, thank you for your immeasurable kindness and patience, your taste in books and TV shows, and for making the times in Oxford incredibly fun. Anna, thank you for your warm and honest friendship, for the wisdom that joy is an art, and for sometimes telling me things I do not want to hear. Anna T, Jemima, and Belinda, thank you for many early sporty mornings, on the river or on the meadow, helping me to keep a clear mind when things got challenging. The usual suspects in Punter evenings and Abbey Road dinners, including, but not limited to, Cecilia, Maria, Nicolò, Felix, and André. Moreover, I am very grateful that coincidence let us live close to each other for a few years, Viola — thank you for always having an open ear and often for your voice of reason.

In Zürich, I have got to know incredibly kind and generous people, who I have found to be able to count on. My colleagues – thank you for making the time here so enjoyable, with after work beers, rowing classes, and raclette evenings. Zsófi, thank you for your loyal friendship, for making me think outside the box, and your relaxed trust in my abilities. Afke, thank you for your trust, limitless understanding, and honesty. András and Eszter, thank you for your warm friendship and support.

Per, thank you for helping me making Zürich feel like home, for your good heart and humour, and, especially in the last months, for your support and for being patient when I could not be.

Last but certainly not least, I want to thank my family – my parents, Elke and Wolfgang, and my brother Sebastian. I would have not got this far without your support, your trust, and your encouragement. Thank you for believing in me, and for making home a place I always look forward to going back to.

Abstract

Social factors have long been included in theories that aim at explaining educational inequality, for example social integration or social influence from significant others. Using social network data from the Children of Immigrants Longitudinal Survey in Four European Countries (*CILS4EU*), I am investigating to what extent social aspects can contribute to our understanding of ethnic and gendered patterns in educational inequality.

The first two empirical chapters focus on explaining ethnic patterns in school grades and in the aspirations to attend university. In these, I find a positive relationship between low school grades and extent of social exclusion, measured through the absence of friendships and the existence of social rejection from classmates. This helps explaining ethnic grade disadvantages of recently arrived migrants, since they are more likely to be socially excluded. Further, I use friendship network data to detect social clusters within school classes, and find that changes in cluster members' aspirations are relatively more important for changes in individual aspirations than the corresponding changes of classmates outside of the social cluster. These chapters use an ego-centric network approach, i.e. they utilise social network data to capture characteristics of the social dimension around individuals and analyse them in regression models on the individual level.

The latter two empirical chapters investigate how social influence can stabilise gendered patterns of favourite subjects and competence beliefs. Examining why girls get discouraged from subjects in the field of Science, Technology, Engineering, and Maths (STEM), I find evidence for influence from friends on favourite subjects, as well as for the tendency of girls to be affected by the preferences of other girls in the classroom specifically when it comes to preferences for STEM subjects. Moreover, I show that there is a social influence from friends on maths competence beliefs, especially for boys, while girls tend to be more influenced by maths grades. These two chapters take a socio-centric approach, i.e. they deploy complete network analysis to detect patterns of social influence, while accounting for network structures and processes.

This thesis shows that social aspects can contribute valuable insights into the study of educational choice and attainment. In identifying concrete social mechanisms surrounding and affecting individuals, this approach can thus help us understand how differences in educational outcomes come about.

Contents

Chapter 1

Introduction and theory

This thesis asks whether social interactions and social structures in the classroom can help explain differences in educational outcomes between boys and girls, and between children of natives and children of immigrants. These social aspects have often been theoretically utilised to explain educational inequality.

In four empirical chapters, this thesis is focusing on ethnic and gendered patterns of school grades and aspirations to attend university, as well as preferences for STEM subjects and maths competence beliefs, respectively. Utilising social network data from the *CILS4EU* data-set, this thesis is investigating both social interactions from the individual perspective as well as the social structures that emerge from those. In that, the analyses capture the social dimension around individuals to understand the extent to which their educational success, in various regards, is embedded in the educational success surrounding them.

This thesis concludes that the study of such social interactions and structures provides valuable insights into the study of individual educational outcomes, and thus, contributes to our understanding of how educational inequality comes about.

1.1 Patterns of educational inequality

Income inequality is often considered to be one of the biggest challenges of our time: Men tend to earn more than women (Gerber and Cheung, 2008), and immigrant groups often experience wage penalties (Dustmann and Frattini, 2012). Factors that crucially determine income and the social position in society are educational choices and success

of the individual: The performance achieved, the type of degree attained, and the field of study.

Educational inequality is understood as the tendency for different sub-groups in society to make different educational choices and be differentially successful in their educational careers. Following the terminology of Charles and Bradley (2002), this can be conceptualised as vertical and horizontal educational segregation. Vertical segregation refers to differences in educational attainment, i.e. the amount of educational stages completed, whereas horizontal segregation refers to stratification in the field of study. These types of segregation can be observed along various different lines, such as gender and immigration background, which are the two factors considered in this thesis. Both the disadvantages of gender and immigration background affect large population groups and are therefore important to be considered; not only for general social equality reasons, but also from a societal, economic perspective. If large population groups fail to reach higher levels of educational attainment or tend not to work in specific sectors, the private and the public sectors miss out on individuals that otherwise could have contributed to a successful economy.

When it comes to differences in the amount of education that men and women achieve on average, i.e. vertical gender segregation, the traditional gender gap of college completion has been reversed in the last decades. Already since the mid-1980s, women's college completion rates began to surpass men's, and today, women earn 58% of all bachelor's degrees in the US. This tendency is mirrored in high school completion rates: more males than females drop out (Snyder et al., 2007). However, horizontal segregation, i.e. the tendency for women and men to choose different fields of study nevertheless contributes to the fact that women on average tend to earn less than men: The field of study also determines income, which especially applies to business, maths, and science-related subjects (Gerber and Cheung, 2008). Thus, horizontal and vertical segregations are related: Equalisation in the horizontal domain can also affect the vertical one, as argued by Charles and Bradley (2002).

Research on ethnic patterns in educational success in European countries have documented divergence, dependent on country of origin and destination: Some ethnic

groups perform rather poorly, especially non-western migrants from less developed countries, or Western immigrants from lower socio-economic backgrounds. This applies for example to youths of Turkish origin in Germany, or to youths of Caribbean origin in the United Kingdom. At the same time, there are some migrant groups that are achieving significantly better results in performance and attainment than natives with the same socio-economic background – these are particularly pupils of East Asian and South Asian origin (Heath et al., 2008). This has been explained by a variety of individual-level factors, but also by the migration history of the country in question, and to positive or negative selectivity of migrants that decide to come to the particular country (Heath and Brinbaum, 2014).

1.2 What explains educational inequality?

As discussed above, educational inequality is the tendency for sub-groups of the general population to differ in their educational attainment or their field of study (vertical and horizontal educational segregation, respectively). Thus, educational inequality is a macro-level phenomenon or observation. There are approaches to explain this that are also on the macro-level, such as the education system, laws, or policies that affect groups in the populations differently (Crul and Vermeulen, 2003; Heath, 2007; Fleischmann and Dronkers, 2010). But, as pointed out by Coleman (1990), causal influence between macro-level phenomena can work only through the micro level. A lot of research on macro-level phenomena, such as educational inequality, thus focuses on individual behaviour and outcomes. That individual behaviour can have drastic effects on the macro level has often been documented, for example in Schelling's influential model of residential segregation (1971), that shows how even weak individual preferences on the micro-level are amplified into complete segregation on the macro level.

A well-known model that links micro-level processes to macro-level outcomes is the "Coleman scheme", or "Coleman boat" (Coleman, 1987, 1990). This can also be applied to explain educational inequality (see Figure 1.1).

The "macro-outcome", i.e. the explanandum, is educational inequality. The "macro-

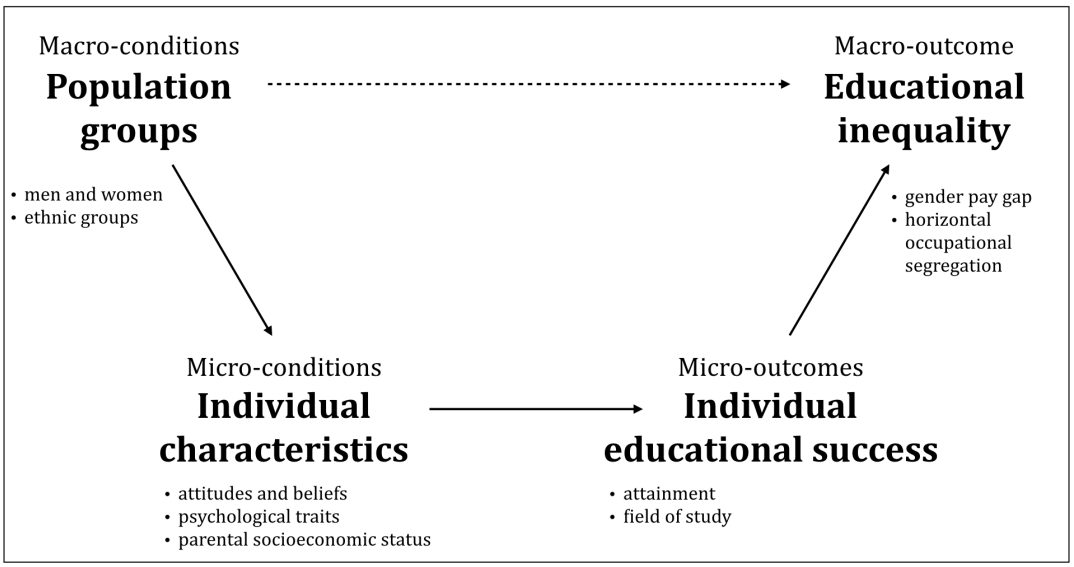

Figure 1.1: Educational inequality explained by Coleman's micro-macro link.

conditions" here are the existence of different groups in society, such as men and women, or different ethnic groups. The micro-level factors are individual characteristics, such as human capital, attitudes, values, or parental socioeconomic status – the micro outcome is individual educational success. Since Coleman's model assumes that macro conditions can only affect macro outcomes through micro-level processes, the assumption here is that gender or ethnicity as such only affect educational success through these individual factors. Furthermore, while it considers macro outcomes to be comprised of events and activities on the micro level, it does not assume that macro-level outcomes are just a simple aggregate of individual outcomes: It assumes a dynamic interplay between individual outcomes, which then affects the macro level. Examples of this are the above-mentioned segregation model by Schelling (1971) or Granovetter's threshold models of collective behaviour (1978). Applying this to the explanandum of educational inequality, this implies that educational inequality as the macro-level outcome is not to be understood as the simple aggregate of individual outcomes, but that there are dynamic processes among individual outcomes that make up the macro level pattern of educational inequality. Thus, this model conceptually allows for interdependencies of individual preferences and choices which shape macro-level outcomes, such as horizontal segregation. An example of this is the tendency for women to focus on specific

fields of study because many other women do so too.

This thesis asks how macro-conditions, i.e. a population stratified by gender and ethnicity, affect individual characteristics or micro conditions, such as aspirations or competence beliefs, and thus individual educational success, the micro-outcomes.

When it comes to an individual's gender and ethnicity, their relation to educational outcomes is not straightforward. In primary school, boys and girls perform similarly well in standardised tests on maths and reading. Throughout secondary school, sex differences emerge and grow larger over time: Girls show an advantage in reading, whereas boys show an advantage in maths. These differences, however, tend to be small (Hyde, 2005), and as pointed out by Spelke (2005), the fact that gender gaps widen over time suggests that social factors are of relevance in this context. There is no reason to assume that this applies differently to immigrant kids, when accounting for skills in the destination country language. Many studies control for this, but still find ethnic penalties in educational outcomes (e.g. Heath et al., 2008; Heath and Brinbaum, 2014).

The link between individual characteristics and individual outcomes has been the focus of much research, especially through the consideration of the socioeconomic background of the family (e.g. Blau and Duncan, 1967; Breen and Jonsson, 2005). There are various ways in which the socioeconomic status of the parents can play a role for educational success. There are direct advantages of ample economic resources for educational success, e.g. to fund tuition fees of universities or private schools, extracurricular activities, or additional tutoring. Furthermore, parents' own experiences in their educational and occupational career can be decisive in what way financial feasibility and necessary effort are perceived by the individual (e.g. Hauser et al., 1983), and can thus affect educational opportunities and choices. Additionally, it has been found that a child's cognitive ability, another important predictor of educational success, is also related to parental background: Children of parents with a high socio-economic status tend to have higher cognitive skills. It is assumed that, besides genetic factors, parenting style and health behaviour are of relevance here (Feinstein, 2003; Ermisch, 2008). Research on social mobility has thus considered the social origin of the family to be a strong

predictor of the level of education and occupational status that an individual attains over the course of their life (Erikson and Goldthorpe, 1992; Breen, 2004). While the socio-economic background of the family has been found to be able to explain much, but not all, of the variation regarding ethnicity, it cannot explain gender differences in educational outcomes. Thus, additional explanations have to be taken into account, such as the social aspects of educational inequality.

1.3 The social aspects

Social resources in general have been found to mediate part of the direct effect of social origin on educational and occupational attainment (e.g. Jencks et al., 1983), and can thus be a counterbalance to disadvantages created by lower parental socioeconomic status. However, access to beneficial social resources is often also compromised by initial inequality in socioeconomic status.

The influential Wisconsin Model of educational and occupational attainment explicitly includes the social dimension (Sewell et al., 1969). Apart from cognitive ability, parental socio-economic status, academic performance, and aspirations, it includes "influence from significant others", such as parents and peers. The relative importance of those two influences largely depends on the individual's age: While in primary school, the parents have the bigger influence, this changes during adolescence, when peers are more influential. Peers are in a similar life situation, and are embedded in the same or in a very similar context. Furthermore, social psychology has found that acceptance and support from peers is substantial for adolescents' well-being, which renders opinions of peers especially determining (Youniss and Smollar, 1985; Osterman, 2000). It has to be underlined, however, that influence is only assumed to be crucial when it comes from *significant others*, i.e. *significant peers*.

It is likely that these significant peers are those few in the classroom that the individual is close friends with. Conceptually, this is based on the assumption that holding up close friendships is only possible with comparatively few people, due to constraints in time and other resources (Baumeister and Leary, 1995). Furthermore, friendship is more likely between people that are similar to each other (McPherson

et al., 2001). This similarity also renders the friend's life situations more similar, and thus the friends' opinions, attitudes, and preferences more applicable to the individual (Lomi et al., 2011).

This influence from significant peers has been studied in two main ways: Through considering the entire classroom context as an approximation of significant peers (e.g. Eitle et al., 2016; Yuan and An, 2017), through explicitly considering the friend characteristics of the individual (e.g. Carbonaro and Workman, 2016; Chen et al., 2017) – or even considering both (Geven et al., 2017). While the investigation of friends is closer to the conceptual meaning of "significant peers", studies aiming at identifying peer *effects* require the explicit inclusion of friendship dynamics themselves in the analysis; this is further discussed below.

Who individuals know, whether they are "significant others" or not, can be a social resource for educational success, and important for behaviour and choices more generally. There are two levels in which social resources can be beneficial for educational success: through individual interactions and through the structure that emerges from and in turn affects those individual interactions. This renders the two domains interdependent, which is in line with general network theory (Carrington et al., 2009). This interdependency is also the reason for the above-mentioned need to account for friendship dynamics when investigating peer, or friend, *effects*.

On the individual level, who individuals interact with can have an impact on educational outcomes through access to beneficial knowledge, or through being exposed and affected by the other's educational goals, attitudes, and behaviour in school. Beneficial knowledge can for instance entail how best to study for a specific test, what kinds of universities to apply to and where to get the best support in that process, and what particular subject to study. This form of social network resource can be conceptualised as quantity and quality of resources that an actor can access through their social network (Lin, 1999, 2000). The importance of access to beneficial knowledge was already discussed by Granovetter (1973), who furthermore concentrated on who the individual is more likely to receive such information from. He argued that acquaintances are a promising source of otherwise inaccessible knowledge, since individuals will be ac-

quainted with more dissimilar people than they are friends with. Those "weak ties" can thus link members of different groups, and thus can provide access to valuable resources.

As mentioned above, the way individuals connect to each other creates structure, which affects opportunity and tendencies of individual interaction as well. This structure in itself can be a resource for educational success, in two main ways. First, if an individual is embedded in a densely connected network (a "closed network"), this can be beneficial, since it implies that everyone is equally connected to everyone else in the network. That entails that information can be accessed directly, and that the possibility of close sanctioning in a community in which everyone knows everyone renders it less risky for people to trust others (Coleman, 1988). Since people tend to connect more to people who are similar to themselves, as mentioned above (McPherson et al., 2001), the more immediate network surrounding is more likely to be similar in terms of resources, too. This can however also be a disadvantage, because it creates information redundancy, and a lack of external impulses. Thus, in addition to closely-knit 'bonding' social capital, 'bridging' capital, consisting of a network tie to other groups that are different in some ways, is likely to provide access to resources which would otherwise be unavailable (Putnam, 2000; Burt, 2007).

In general, inequality in social resources, on the individual as well as the structural level, can lead to cumulative advantage, also known as the Matthew effect (Merton, 1968): Those who are already at an advantageous position will be able to further this more due to access to the social resources that this network position carries. For example, it has been found that job seekers of initially higher socio-economic status are able to find more, and more high-quality jobs than those of lower socio-economic status (Ioannides and Datcher Loury, 2004), and that men can exploit those kinds of resources better than women (Åberg and Hedström, 2001).

This social dimension of explaining educational inequality is thus a complex level that explains how individual conditions lead to individual outcomes, through social interactions and social structures. It is useful to apply the Coleman scheme to this relationship as well, as described in the next paragraph.

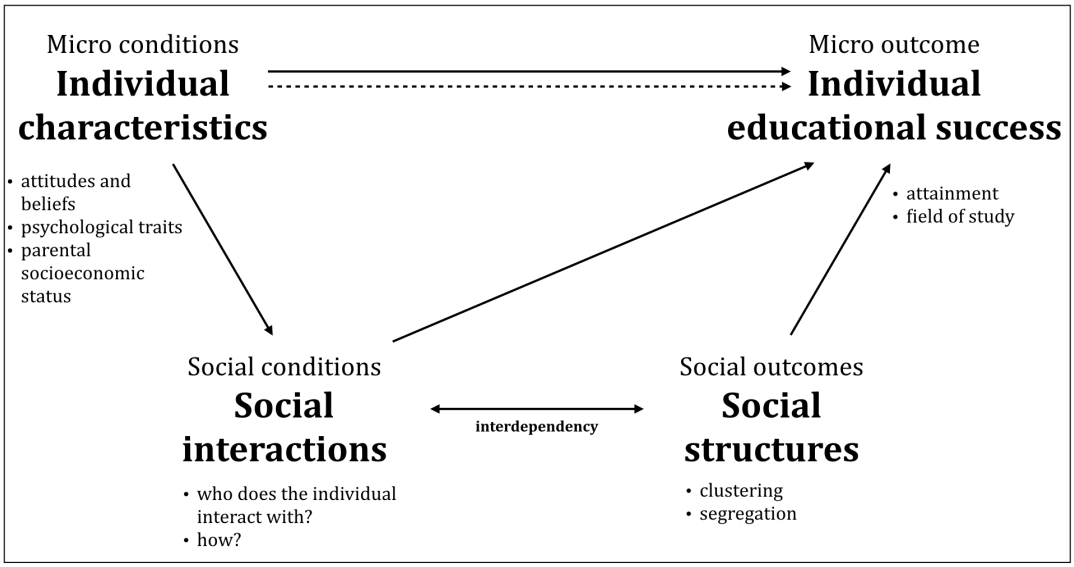

Figure 1.2: The "Micro and Social" Coleman boat of educational inequality.

1.4 A multilevel Coleman boat of educational inequality

This thesis proposes to integrate the social aspects of educational inequality, as discussed in the previous section, into a Coleman scheme of educational inequality. Figures 1.2 and 1.3 show the conceptual framework of this thesis. Figure 1.2 depicts how individual characteristics are negotiated the social sphere, and how this affects individual outcomes; Figure 1.3 integrates this into Figure 1.1, which results in a "multilevel Coleman boat".

The extension to the Coleman scheme that this thesis is proposing and exploring is aiming at further explaining the link between individual characteristics and individual educational success. While there are some ways in which these two link directly, they are not able to explain the entire relationship, as discussed above: Social aspects additionally link them in a way that can be considered as parallel to the way in which micro processes link macro phenomena (see Figure Figure 1.3)

Following the general idea of the Coleman scheme, individual characteristics affect social interactions, leading to different tendencies in how individuals interact with others. As argued above, the social interactions and social structures are interdependent: Consequently, the social conditions and social outcomes both affect individual educational outcomes, unlike in the classic Coleman scheme, in which only micro outcomes,

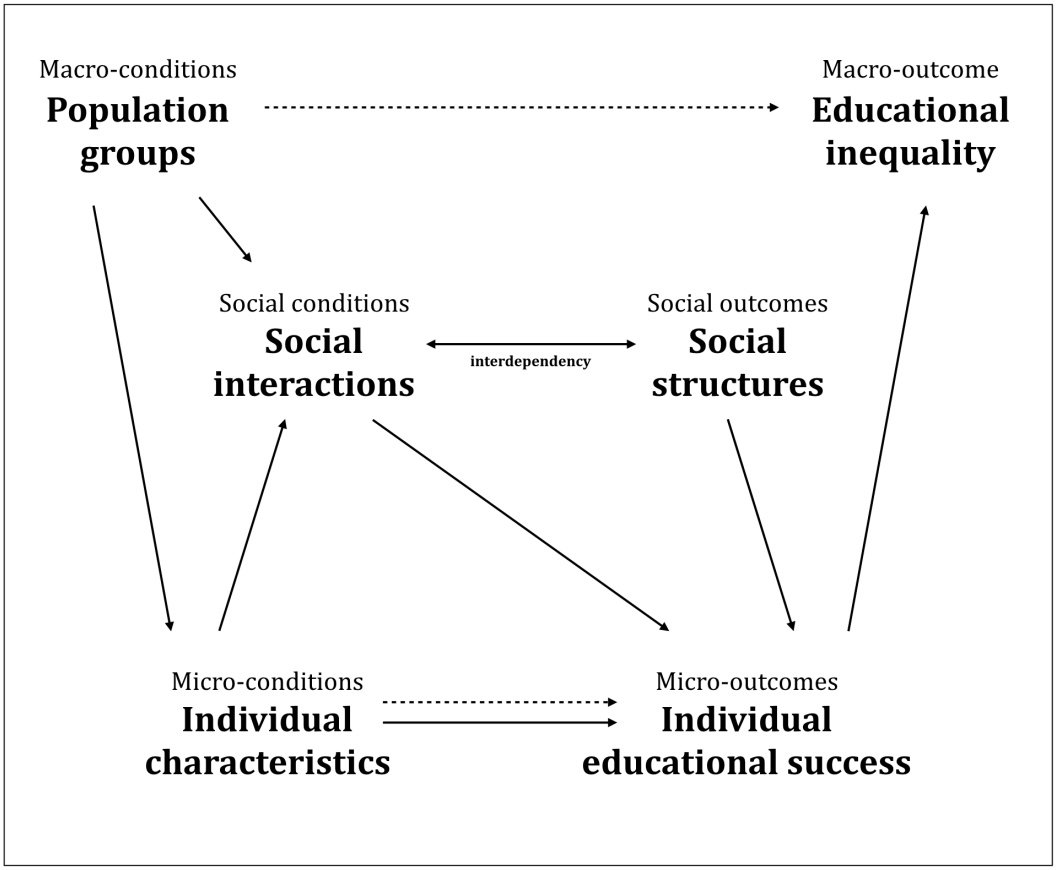

Figure 1.3: The "Multilevel" Coleman boat of educational inequality.

and not micro conditions, affect macro outcomes.

The social domain, consisting of the interdependent social interactions and social structures, can be studied in a number of ways. First, there is the ego-centric approach, which considers the specific interaction types that an individual is involved in, but not the way people that the individual interacts with are connected to each other. As discussed above, these types of analyses can give valuable insights, however, they cannot make any claims about effects, since they cannot account for endogenous friendship network dynamics. They can, however, show correlations and document tendencies on a level that is comparably very exact. As discussed, a lot of research uses aggregates of individuals' characteristics on the classroom level: Using friends instead of everyone in the class is thus a finer measurement.

Second, there is the socio-centric approach that takes into account individuals' interactions as well as the structure that emerges from and interacts with them, such

as segregation, hierarchy, or density. When taking this approach over time, it can consider network dynamics as well as individual outcomes, and thus explicitly consider peer effects. This is due first to general homophily in friendships, and second to the endogenous nature of friendship network processes. First, people tend to become friends with people with whom they share many or important attributes (such as age, gender, attitudes, preferences in cultural consumption or free time activities), a phenomenon known as homophily (McPherson et al., 2001). Separating out friend similarity due to such selective friendship from similarity as a consequence of peer effects requires a longitudinal approach, and is an integral part of this thesis. Second, who becomes or stays friends with whom is not only influenced by attributes: network processes like reciprocating friendships (reciprocity), becoming friends with friends of friends (transitivity), or popularity of individuals in the network are just a few examples. The way these processes shape a network is endogenous, i.e. they depend on each other: An individual is for example more likely to reciprocate a friendship nomination by someone who is also a friend of a friend. Since this interdependency is violating the basic assumption of independence of regression analyses, social network analysis is needed to account for these network processes (Carrington et al., 2009). Since a longitudinal approach is needed to separate selection from influence effects, as discussed above, longitudinal social network analysis is needed when aiming at effects from friends (Steglich et al., 2010), as demonstrated in a number of studies (e.g. Schaefer, 2016; Lakon et al., 2017).

The chapters in this thesis makes use of both the ego-centric and the socio-centric approach of studying the link between individual characteristics, gender and migration background specifically, and individual educational success. In that, they are all taking a unique approach to answer the overarching research question of this thesis, i.e. whether social factors can help explain differences in individual educational outcomes, to eventually understand educational inequality. Chapters 2 and 3 are taking an ego-centric approach and focus on differential educational outcomes of immigrants. That means that they study social interactions from an individual perspective, to understand the way individuals are embedded in the social structure around them, without

accounting for endogenous processes and the structure itself. This is useful to get an initial understanding how how individual educational success is associated with the success around individuals; furthermore, it is methodologically more straightforward and accessible.

Chapters 4 and 5 use the socio-centric approach to study gender differences, i.e. they are able to not only account for the immediate interaction structure surrounding the individual, but also for the structure and endogenous processes that emerges from and in turn affects social interactions. While this is computationally much more advanced and elaborate, this approach also has a lot more power and is able to accurately map concrete social mechanisms and their effects on individuals over time.

In deploying both approaches in two empirical chapters each, this thesis unites the advantages of both to investigate the extent to which the social environment of individuals does play a part in determining the educational success of the individual.

1.5 Country contexts

This thesis is using the *CILS4EU* data from four European countries: England, Germany, the Netherlands, and Sweden. Chapters 4 and 5 use the Swedish part of the *CILS4EU* data only, due to technical reasons; this is further discussed below and in the methods parts of the respective chapters. Chapters 2 and 3 use pooled data from all countries, and find some differences between countries. These chapters are concerned with differences for different immigrant groups, which is why most cross-country differences are likely to be explained with the respective immigration history and the status of the different ethnic groups in those countries.

Immigrant groups in England have been shown to be most favourably selected, with ethnic minority groups outperforming the majority group, with a long history of immigration from its current or former oversea territories and associated long period of cultural and linguistic exposure (Heath et al., 2008). In Germany, the Netherlands, and in Sweden, after histories of guest worker programmes, colonial involvement, and free Nordic labour migration, respectively, immigration related to humanitarian reasons has increased immigration rates in the past decades, particularly so in Sweden.

All three countries have received migrants from the Balkans, the Middle East, and Northern Africa in the past years (Dustmann and Frattini, 2012), although the most recent migration from Syria took place after the collection of the data in 2010. We can thus expect to see more ethnic differences when it comes to newly arrived immigrants in Germany, the Netherlands and Sweden. Additionally, for Germany, more persistent disadvantages in the older generations are also likely to exist, due to ethnic segregation resulting from guest worker programmes (Musterd, 2005). These differences in ethnic disadvantages have implications for the extent to which social interactions are able to explain part of them: Where migrants are outperforming natives, such as in the case of Indian or Chinese students in England (Heath et al., 2008), it is possible that social exclusion is part of a different social mechanism. Furthermore, resulting differences in group sizes have implications for interactions and structure, due to ethnic segregation. This will depend on host country, migration generation, as well as ethnic group, and thus cannot be discussed in depth here. One example is the persistent disadvantage of immigrants with a Turkish background in Germany, since the majority of this population migrated to Germany as part of a guest worker programme. These programmes implied that the population lived and worked very much segregated from the German majority.

Furthermore, there are differences in the education systems of the four countries. Germany and the Netherlands have a tracked education system, whereas Sweden and England do not. For performance, as investigated in Chapter 2, this is not likely to play a huge role, but it is for the results of Chapter 3, which consider university aspirations. Since only graduating from the higher tracks enables students to attend university, and tracking in Germany and the Netherlands has already happened at the time at which adolescents are interviewed, students in these countries can reasonably estimate the probability that they will attend higher education. This insight comes later in Sweden, which is the third country considered in Chapter 3.

Chapter 2

Social exclusion and school achievement

Ethnic minorities in Europe show diverging patterns of educational success, but in most national contexts, migrants and children of immigrants have lower achievements in terms of grades than their majority peers. This study asks whether social exclusion in the classroom can contribute to explaining this pattern. While limited access to social resources is often assumed to be of significance for educational success, it has rarely been measured explicitly. In this study, social exclusion is measured accurately and on a large, cross-national scale, by using social network data from 731 classrooms in England, Germany, and Sweden (from the *CILS4EU* data). Results show that social exclusion is negatively associated with school grades, but this does not contribute much to understanding grade differences between children of immigrant and majority background.

This chapter is based on the paper "Social exclusion and school achievement: Children of immigrants and children of natives in three European countries", that has recently been accepted for publication in *Child Indicators Research*.

2.1 Introduction

Social resources are crucial for educational success: Social exclusion in the classroom can severely compromise access to those resources, which can have detrimental effects on educational success and thus, in the long run, socioeconomic attainment. This applies in particular to groups that are vulnerable to begin with, such as children of immigrants. Strategic knowledge about the education system, about manners, values and cultural codes that prevail at school, and access to networks of support and inspiration that promote learning and performance are typically less available to children of immigrants than to their ethnic majority peers. Their parents, migrants themselves, are more likely to lack these types of resources, rendering the existence and quality of social relations and connections at school more critical and decisive for their educational career. If ethnic minority children are able to establish a network of support and resources for themselves within school, this can counteract those disadvantages. However, if ethnic minority children are socially isolated and avoided, these types of support networks are less likely to be established. This inequality of network resources then contributes to the perpetuation of already existing ethnic stratification patterns. Thus, the main question that this paper seeks to answer is: Can social exclusion help explain ethnic differences in school achievement?

While it is not a new idea to consider social network resources (especially Lin, 1999, 2000), the fact that social exclusion is difficult to measure adequately on a large-scale basis has resulted in a scarcity of empirical research on its consequences. Through the utilisation of social network data, this paper adequately measures social exclusion in the form of social isolation (absence of friendship) and social avoidance (negative relationships). Using the *CILS4EU* data (Children of Immigrants Longitudinal Survey in Four European Countries), the analysis is furthermore able to test this dynamic in several national contexts. This study uses sociometric data, i.e. social network data, on 15,017 students in 731 classrooms in England, Germany, and Sweden. These countries differ in their institutional arrangements and patterns of ethnic educational stratification, and the variation in contexts is particularly valuable for testing the occurrence of general mechanisms. The high number of networks and the possibility of generalizing

the results to populations in several countries are exceptional due to high costs and effort associated to the collection of social network data.

2.2 Theory, previous research, and hypotheses

This section gives an overview of relevant theory and previous research, and develops the hypotheses guiding the research presented in this chapter.

2.2.1 The social dimension of educational success

Research in social stratification has long focused on diverging educational success, both attainment and performance, with the traditional predictor being the socioeconomic background of the family. Following increased immigration rates in many Western societies, recent research has increasingly considered ethnic inequalities in educational success (e.g Heath et al., 2008). While research on first generation immigrants has singled out language skills and pre-migration educational qualifications as obstacles to achieve higher education (van Tubergen and van de Werfhorst, 2007), a lot of recent research on ethnic inequalities has focused on the second generation, that is, children born in host countries to immigrant parents, where other explanations need to be sought. These results show that the socio-economic origin of immigrant groups is behind much, but not all, of the overall poorer performance of ethnic minority groups (Heath and Brinbaum, 2014). A potential explanation for the remaining minority disadvantage is the social dimension of youths, as expressed in their social networks, friendship patterns, and social exclusion such as bullying and social isolation. Such a focus acknowledges that the role of peers becomes more significant during adolescence (Youniss and Smollar, 1985; Osterman, 2000). Since adolescents spend the majority of their time at school, the school class is a particularly significant social context for them, and as a societal 'micro-system' the school class harbours both positive and negative peer relations. The aim of this paper is to analyse whether negative relations may lead to disadvantage in school achievement, and whether this more severe to the ethnic minority group.

2.2.2 Social exclusion and limited access to social network resources

When it comes to negative social relations at school, most research has focused on bullying and its negative effects on a variety of outcomes (e.g. McDougall and Vaillancourt, 2015), however, subtler forms of negative relationships are likely to play a role, too. In particular, isolation, i.e. having no friends, and being avoided by others are forms of a non-positive social environment in the class, which have not been researched as extensively as bullying. Both forms are less explicit than bullying, but can nevertheless be detrimental when it comes to individual outcomes, including educational success.

There are some factors suggesting that ethnic minorities are at a higher risk of being socially excluded. Social misfit theory proposes that individuals are avoided when they are different in one or several regards (Wright et al., 1986), which could be the ethnic or cultural background. Furthermore, recent studies find that youths with a migration background in Sweden are not only more likely to be victimised (e.g. Vervoort et al., 2008), but are also at a higher risk to face subtler forms of exclusion (Plenty and Jonsson, 2017). Work on friendships and friendship networks, has shown repeatedly that friendship is more likely and more stable between individuals that are similar to each other, for example regarding their ethnic background, a phenomenon known as homophily (McPherson et al., 2001). Thus, depending on the ethnic composition in the classroom, ethnic minorities can be at a higher risk of not being able to form or initiate less stable friendships if there is not much opportunity of befriending someone with a similar ethnic background.

While the need to belong is a fundamental human need that applies to people of all ages (Baumeister and Leary, 1995), it has been found that belonging and social acceptance is particularly important for adolescents' well-being (Brown, 2004), relative to adults or children. Previous research has documented the negative effects of low mental well-being on educational outcomes (Nakamoto and Schwartz, 2010), but there can also be direct effects from being isolated and avoided on academic achievement, since social exclusion will reduce the chance of learning from others, doing homework together, or drawing on friends' parents for academic support, such as strategic knowledge on how to navigate the education system (Erikson and Jonsson, 1996). These social network

resources can be conceptualised as quantity and quality of resources that an actor can access through their location in a social network, i.e. through the people he or she knows, and who these people know (Lin, 2000). These types of resources have been found to mediate part of the effect of social origin on educational and occupational attainment (e.g. Jencks et al., 1983), and may counterbalance disadvantages created by lower parental socioeconomic resources. However, access to beneficial social resources is often also compromised by initial inequality, which can aggravate these already existing inequalities. Since personal networks tend to be segregated along a variety of personal characteristics, such as gender and age, but also socioeconomic status (McPherson et al., 2001), there is an increased risk for initial inequalities to be reinforced: Those who are already at an advantageous position will be able to promote this more due to access to the social resources that this network position carries, and those who are already disadvantaged cannot exploit their social contacts to the same extent, and are left behind even more (Lin, 2000).

2.2.3 Isolation and avoidance in the classroom

Classrooms, in which children and adolescents spend a lot of time, provide a potential opportunity for gaining access to social network resources. Although classrooms tend to be sorted by ability and parental background, due to residential segregation or in some countries due to a tracked education system, they are still an important environment for associating with others in the same life situation but with a potentially different background: ethnic as well as socioeconomic. Therefore, school classes have the potential to provide a platform for the access of valuable social network resources, but social exclusion can inhibit the access to such resources.

There are several different ways in which pupils can be socially excluded, such as isolation and avoidance. Isolation here is understood as the absence of personal links (e.g. friendship ties), and can either be understood with reference to everyone in the classroom (complete isolation), or just with reference to the majority students in the classroom (isolation from majority students). Avoidance implies negative relations, such as dislike. These two types of social exclusion are likely to affect grade point

average through slightly different mechanisms, i.e. lack of support and limits in access to information due to social isolation, and in addition stress and negative mental health through avoidance. While this analysis is not aiming at identifying the specifics of how social exclusion can affect grade point average, the fact that these two types of exclusion are likely working through different mechanisms warrants the differentiation and separate consideration of isolation and avoidance.

Social resources of students in classrooms in which they are isolated and/or avoided are compromised, thus, drawing on the discussions above, the first hypothesis states:

Hypothesis 1: Social exclusion has a negative effect on educational performance.

While ethnic minority pupils are expected to be more socially excluded, it is furthermore assumed that this varies with respect to the generational status, i.e. how long individuals' families have been in the host country. Especially newly arrived migrants, i.e. those that have only been in the country for a couple of years, are at a higher risk of being socially excluded, for example due to language barriers or simply because they have had a shorter time to make friends. The second generation, who were born in the host country, will less frequently face such problems, but may still not be fully integrated socially due to homophily tendencies as well as prejudice. In combination with hypothesis one, it is thus hypothesised:

Hypothesis 2: The negative effect of social exclusion accounts for immigrant disadvantages in educational success.

Hypothesis 3: This tendency is expected to be strongest for those who migrated themselves.

Furthermore, social exclusion may have more detrimental consequences for school achievement for children of immigrant origin than for those with native-born parents. This is because the latter have gone through the host-country educational system them-

selves, and can thus offer better advice than parents who have never attended school in the destination country. In addition to this practical help, language barriers could render it more difficult for parents of students with a migration background to seek information on opportunities and support that might be available. Thus, children of immigrant origin are likely to be in more need of social resources outside the family. Consequently, the fourth hypothesis states:

Hypothesis 4: The negative association between social exclusion and educational performance is stronger for students with a migration background.

With the utilisation of social network data to measure social exclusion on a large scale, this analysis manages to explicitly test a mechanism that has received a lot of attention in theoretical works, but which has not received much empirical attention, due to lack of suitable data.

2.3 Data and Methods

This section introduces the data and methods used in the analysis.

2.3.1 Data

The data analysed in this paper is from the first two waves of the Children of Immigrants Longitudinal Survey in Four European Countries (*CILS4EU*) data-set, which focuses on the integration of first and second generation youths in four European countries: England, Sweden, Germany, and the Netherlands (Kalter et al., 2014, 2015). This is the first comprehensive dataset that is fully standardised, longitudinal, and generalisable to the survey country populations, providing data on youths, teachers, and parents, as well as information on class-room social networks. Wave 1 of this survey interviewed students at age 14-15, and wave 2 at age 15-16. The survey includes information on minority and majority ethnic groups, oversampling immigrant-dense schools in order to increase power in the analyses that focus on ethnic minority students.

This study uses data from England, Germany, and Sweden. Since the education

system in the Netherlands is not strictly classroom-based to the same extent as in the other countries, the measure of social exclusion through the classroom-networks is regarded to be less valid than in the other countries; thus, this country is excluded from the analysis. In total, the sample used here has information on 15,017 students in 731 classrooms, in 381 schools. Mostly, two classrooms per school were included in the sample, with an average of 20.6 students in each classroom.

2.3.2 Variables

Descriptives of all variables used are reported in Table 2.1, in the full data set and in the analytical sample, respectively. Around one third of all respondents had missing information on one or more variables that were included in the analysis, mainly due to missingness on the outcome variable, grade point average (GPA; 3,328 of 15,017 cases). To deal with item non-response on the independent variables, missing values were imputed using multiple imputation with 10 imputed data-sets. This was done for variables that are not based on network survey items, since imputing network ties is not straightforward and currently only under development. Results based on non-imputed data can be found in Tables A.2 and A.3 in the Appendix. Analyses based on imputed data and list-wise deletion yield the same substantial conclusions, with only little variations in effect sizes and standard errors.

The outcome variable GPA was collected in wave 2, while the predictor variables are taken from wave 1.

Outcome variable: grade point average (GPA)

The outcome is GPA, measured as the average of the school grades in mathematics and the survey country language. This is available as self-reported grades for England and Germany, and as teacher-assigned grades taken from school-register data for Sweden. This renders the data for Sweden more objective, and presumably of higher quality. The measure for GPA is z-standardised within each country, because countries use different grading scales.

Table 2.1: Variable and sample descriptives.

	Full sample						*Analytical sample*					
	Mean	SD	Median	Min	Max	N	Mean	SD	Median	Min	Max	N
Grade point average (GPA)	0.00	1.00	-0.08	-4.38	2.67	11689	0.09	0.98	-0.08	-4.38	2.67	8376
Age (in months)	180.01	6.99	179.00	157.00	228.00	14256	179.59	6.87	179.00	157.00	219.00	8376
Cognitive ability index	18.34	4.48	19.00	0.00	27.00	13982	18.92	4.18	19.00	0.00	27.00	8376
Highest parental degree	2.92	0.96	3.00	1.00	4.00	13797	2.98	0.92	3.00	1.00	4.00	8376
Parental unemployment	0.25	0.43	0.00	0.00	1.00	12230	0.23	0.42	0.00	0.00	1.00	8376
Other language spoken at home	0.39	0.49	0.00	0.00	1.00	14185	0.36	0.48	0.00	0.00	1.00	8376
School resistance	8.78	3.65	8.00	0.00	25.00	15017	8.83	2.90	8.00	0.00	25.00	8376
Isolated from everyone	0.03	0.16	0.00	0.00	1.00	15017	0.00	0.07	0.00	0.00	1.00	8376
Isolated from natives	0.14	0.35	0.00	0.00	1.00	15017	0.11	0.31	0.00	0.00	1.00	8376
Avoidance by majority	1.32	2.11	1.00	0.00	20.00	15017	1.30	2.08	1.00	0.00	20.00	8376
Avoidance by minority	1.04	1.23	1.00	0.00	17.00	15017	1.00	1.15	1.00	0.00	17.00	8376
Number of friends (indegree)	3.19	2.05	3.00	0.00	18.00	15017	3.46	2.00	3.00	0.00	14.00	8376
Country (all)						15017						8376
England						4442						2192
Germany						5127						3137
Sweden						5448						3047
Gender (all)						14344						8376
Male						7256						4192
Female						7088						4184
Generational status (all)						14308						8376
Native						8963						5578
Newly arrived						526						211
First generation						1276						605
Second generation						3543						1982

Main predictor 1: Migration background as generational status

Migration background is included in the models as a categorical variable, based on the official generation status variable available in the *CILS4EU* data-set, constructed according to the "ancestral distance" of the respondents to their immigrating ancestors (Dollmann et al., 2014). The official variable was constructed using information on country of origin and age of arrival in the survey country of the child, his or her parents as well as grandparents. Based on this, the variable takes four values for the purpose of this analysis. First, "majority" refers to adolescents whose parents were born in the survey country. Second, "newly arrived" captures those that arrived after the age of 10, who have been in the country for only a few years: at the point of time of the survey, youths were mostly between 14 and 15 years old. Third, the "first generation", i.e. those who migrated themselves, but arrived before the age of 10. Fourth, the "second generation" refers to those who were born in the survey country, but whose parents have been born abroad. Since the main analysis here pools data from the three countries, differentiating between countries or regions of origin is not feasible, as the survey countries vary substantially in the ethnic composition of their immigrant populations; this would have led to a low number of cases in some categories. Thus, this paper utilises the generation status in the main analysis, and uses the ethnic composition of the immigration population in the country-specific interpretation of the results. The ethnic composition of the generational status categories of the pooled data is presented in Table A.1 in the Appendix. Additionally, country-specific analyses are included in Tables A.4 to A.9 in the Appendix.

Main predictor 2: Social exclusion

In line which what has been discussed in the literature review, social exclusion is conceptualised here both as social isolation and social avoidance, by accounting more closely for the type of isolation, as well as for the migration background of those who avoid the individual.

1. Isolation from everyone in the class. This dichotomised variable captures social isolation as the complete absence of friendship nominations by classmates. 3.6%

of majority students and 3.8% of minority students are isolated from everyone in the class.

2. Isolation from majority students in the class. This dichotomised variable accounts for individuals not receiving friendship nominations from majority students in the classroom, but from minority students. They are thus not isolated completely, as described in point 1, but they are isolated from the majority students in the class. 3.5% of majority students and 31.1% of minority students are isolated from majority students in the class, and have no majority friends outside the classroom, either.

3. Avoidance by majority students in the class. This variable captures the number of incoming avoidance ties from majority students in the classroom, and expresses this as a proportion of all majority students in the classroom. The avoidance tie is based on the survey question "Who do you not want to sit next to?". A value of 0.5 would thus mean that 50% of majority students in the classroom report that they do not want to sit next to the individual.

4. Avoidance by minority students in the class. Following the same logic as described in point 3, this variable measures avoidance by minority-students in the classroom.

For measures 3 and 4, avoidance ties from those that the individual considers to be their friend are not counted. Among friends, there can be various reasons of why they do not want to sit together, such as being able to better concentrate on the lessons. Approximately 6% of 37,532 recorded avoidance ties (across all three countries) were dropped based on this reasoning. Additionally, the analyses were carried out without dropping those ties. The results suggest that the measure of social avoidance is more accurately tapping into the dimension of interest here if the ties from friends are dropped (more details are provided in the section on robustness checks).

Figures 2.1 and 2.2 show how the measures for avoidance and isolation apply by migration background. These descriptives show two main aspects. First, minority students are a lot more likely to be isolated from majority students than majority

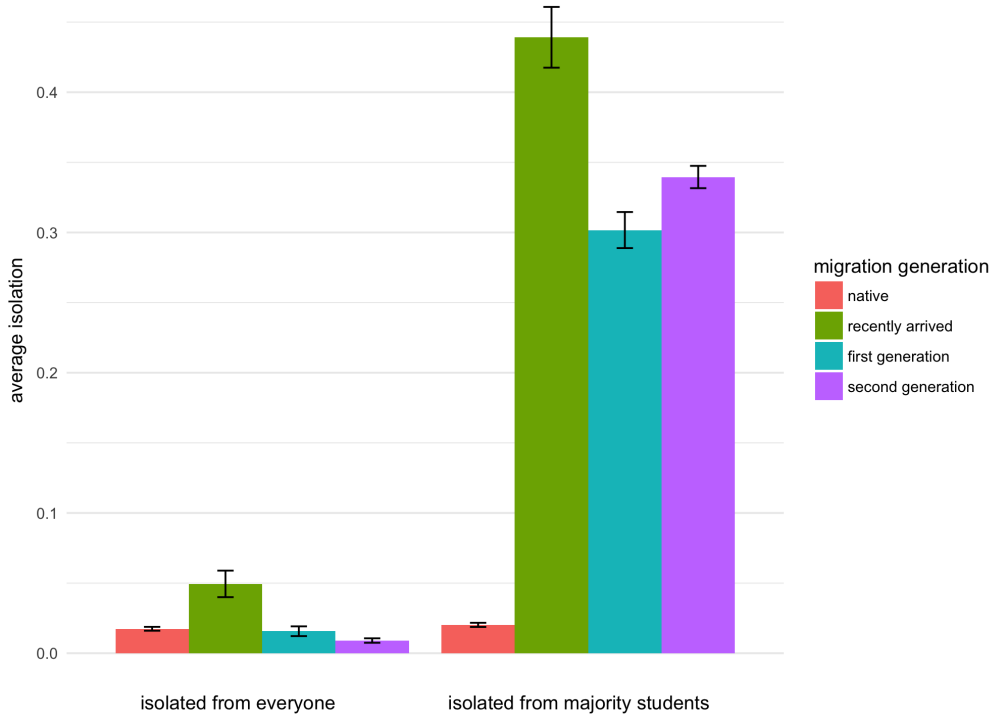

Figure 2.1: Isolation by migration background

students themselves are; this applies in particular to newly arrived immigrants. Second, avoidance ties seem to exist more within groups, i.e. majority and minority, than between groups. Newly arrived minority students are more often than other minority students avoided by majority students. This supports the above-discussed tendency of newly arrived immigrants being particularly vulnerable to social exclusion, especially to social exclusion from majority students.

Controls

The analysis furthermore controls for a variety of individual characteristics which are likely to affect grade point average, too. Besides gender and age (in months at the survey date), the models include a measure of cognitive ability. This is the result of a language-free test that is part of the *CILS4EU* data, and which is based on solving graphical problems in a given time (CILS4EU, 2014). This is included in the model to control for baseline differences in ability. Predicting grades controlling for cognitive ability accounts for the individual's potential to get high grades at school, which is an

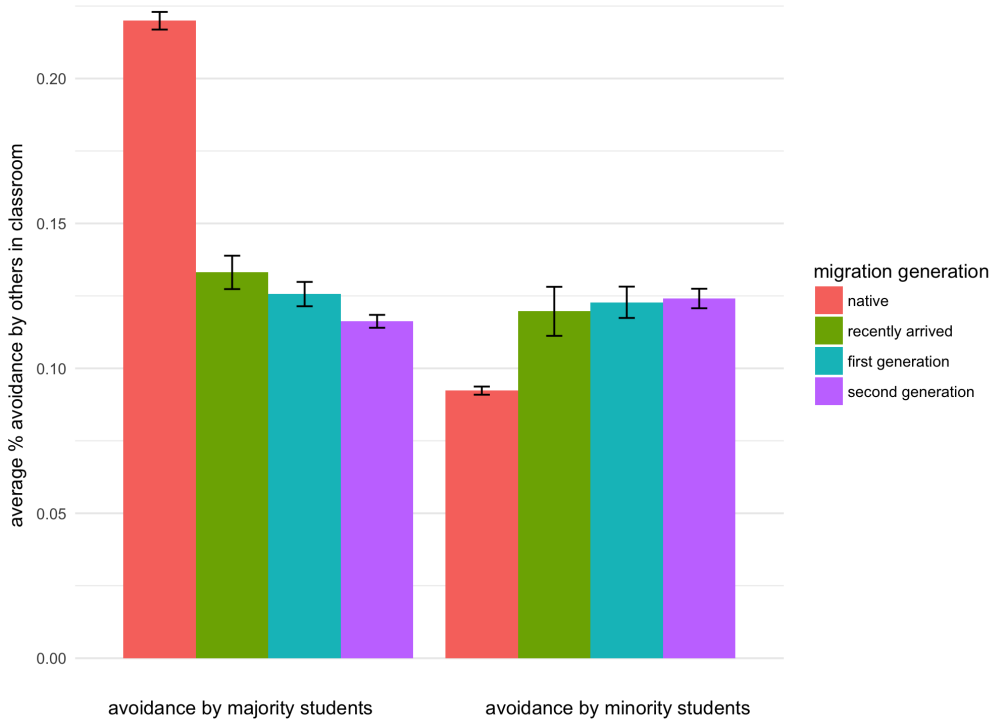

Figure 2.2: Avoidance by migration background

important, but not perfect predictor.

To account for baseline differences in language ability, which is crucial to include when considering ethnic differences, a dummy of whether another language, besides the survey country language, is spoken at home. This is preferred to a language ability test, which is included in the data as well, because of potential endogeneity in this test variable (those of immigrant origin probably improve their language proficiency if they have majority friends). Additionally, the language test result is highly correlated with the cognitive ability index. Parental background is controlled for by the highest level of parental educational attainment, which includes four categories: primary school, lower secondary school, higher secondary school, and university degree. A control for parental unemployment is included in the models as well, since past research has shown that youths with fewer economic resources are more likely to be excluded from activities and friendships (Hjalmarsson and Mood, 2015). The unemployment variable is a dummy, accounting for whether there is not at least one parent living with the adolescent who is employed.

Resistance to schooling is controlled for, too, as it is an important alternative explanation for low grades. Following Geven et al. (2017), a measure for resistance to schooling is computed based on five survey items that capture the extent of the individual (1) arguing with teachers, (2) getting punishments at school, (3) skipping classes, (4) being late to classes, and (5) putting effort into work for school. These items load on one factor with individual loadings between 0.4 and 0.8 (Crohnbach's alpha = 0.71). The items were combined additively, with the fifth item being reversed; non-complete cases on all five items were assigned a missing value on the scale. Furthermore, the models control for number of friends, measured as the in-degree, i.e. the number of people who nominate the individual as a friend. This control is included here since the aim of the analysis is to understand the effect of other types of social exclusion besides the effect of absence of friends. Three of the four measures used in this analysis do not exclude the possibility of having friends. Those who are considered to be isolated from majority students only, have friends who do not belong to the majority group. The two measures of avoidance do not consider friends at all, and only account for avoidance relationships. The effect of isolation from majority students or avoidance net of number of friends can show if it is the exclusion from a certain sub-group of people in the classroom that makes a difference, rather than being completely excluded from everyone.

To deal with item non-response, missing values were imputed using multiple imputation with 10 imputed data-sets. This was done for variables that are not based on network survey items, since imputing network ties is not straightforward. The dependent variable was not imputed.

Official survey weights were used in all models to account for the over-sampling of immigrant-dense classrooms.

2.3.3 Methods

Linear regression models with standard errors clustered at classroom level are utilised to predict the grade point average (GPA) from *CILS4EU* wave 2; predictors are from wave 1, with approximately one year between the two waves. After confirming that

there is no country-heterogeneity in the social exclusion effects, which was tested by including country interactions (see Table A.10 in the Appendix), analyses were run on pooled data from the three countries. Additionally, all models were run separately for each country (see Tables A.4 to A.9 in the Appendix): where country-specific differences are found, they are discussed in the results section. In the first set of models, Model 1 predicts GPA only by migration background. In the next step, the controls described above are added (Model 2), then measures of social exclusion (Model 3). In the second set of models, interaction effects between migration background and all social exclusion variables are added to the full model configuration. This is done with a dummy of migration background (native, i.e. belonging to the majority, vs. first or second generation) and separately for each measure of social exclusion.

All analyses have been carried out in *R*, version 3.2.2, using the *plm* package for the regression analysis, and the computation of clustered standard errors. Clustered standard errors are calculated according to Stock and Watson (2008), which is the standard procedure in many statistical packages, such as *Stata*.

2.4 Results

Before the results of the step-wise regression analyses are discussed, the relationship of interest here, i.e. between social exclusion and GPA, is explored descriptively.

2.4.1 Descriptive results

Figure 2.3 shows the relationship between GPA and being socially isolated from everyone, and from majority students, respectively. Figure 2.4 shows the relationship between GPA and being avoided by majority students, and by ethnic minority students, respectively. For the purpose of this graph, an individual was considered to be socially avoided when at least 30% of majority children in the classroom (or ethnic minority students, respectively) reported that they avoid the individual – this was considered to be a high enough, yet still somewhat conservative, threshold. All associations between the social exclusion measures and GPA are presented separately for majority students and ethnic minority students.

When comparing the type of social exclusion, i.e. isolation (Figure 2.3) and avoidance (Figure 2.4), there are some similar patterns, and some differences. First, socially excluded pupils (either of minority or majority origin) get lower grades than non-excluded students. Second, non-excluded majority children get the highest grades, while excluded ethnic minority students receive the lowest grades, in all four cases. There are, however, also some differences between isolation and avoidance. First, majority children who are isolated from everyone tend to get significantly lower grades than those who are avoided. Second, ethnic minority students who are avoided by majority students tend to get significantly lower grades than those who are isolated.

These descriptive results point to the existence of the relationship hypothesised in this paper, i.e. a negative association between grades and being socially excluded, as well as to a tendency for ethnic minority students to be more vulnerable when they are excluded.

2.4.2 Results from regression analyses

Hypothesis 1, stating that social exclusion has a negative effect on educational performance, is supported by the analyses in Table 2.2[1]. Of the four exclusion measures, two have a significantly negative effect on grade point average: Being isolated from majority students and being socially avoided by majority students in the classroom. This applies while controlling for ethnic background, socio-demographic, cognitive and personality-related factors. Notably, this also applies when controlling for the number of in-going friendship nominations, i.e. the number of people reporting that they are friends with the individual. This points to the fact that social avoidance, especially by majority students, has a negative effect, even if the individual has other friends in the classroom. The net effects, however, are relatively weak: Being isolated from majority students is associated with a 0.09 standard deviation's reduction in grades across our three countries, and a similar reduction in avoidance by majority students by 25 percentage points is related to a similar grade reduction (0.39/4=0.10).

While the direction of the effect is the same in all three countries, there are some

[1]Results based on non-imputed data are shown in Table A.2 in the Appendix

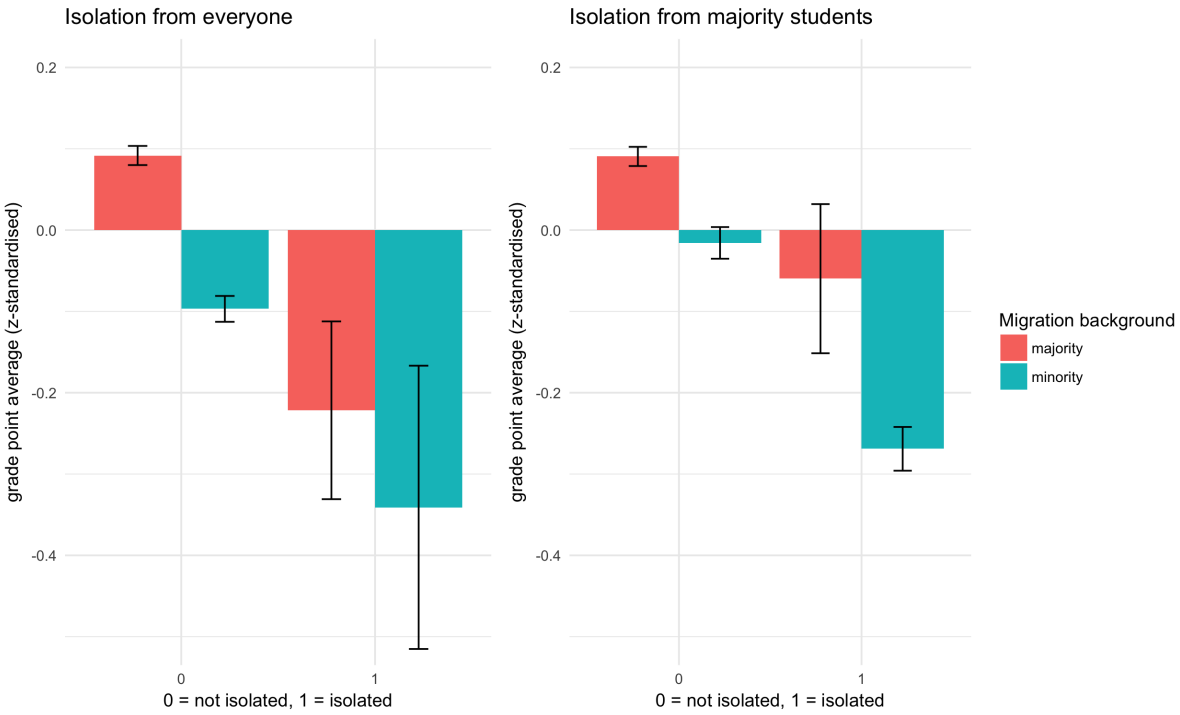

Figure 2.3: Grade point average by extent of social exclusion (isolation) and migration background.

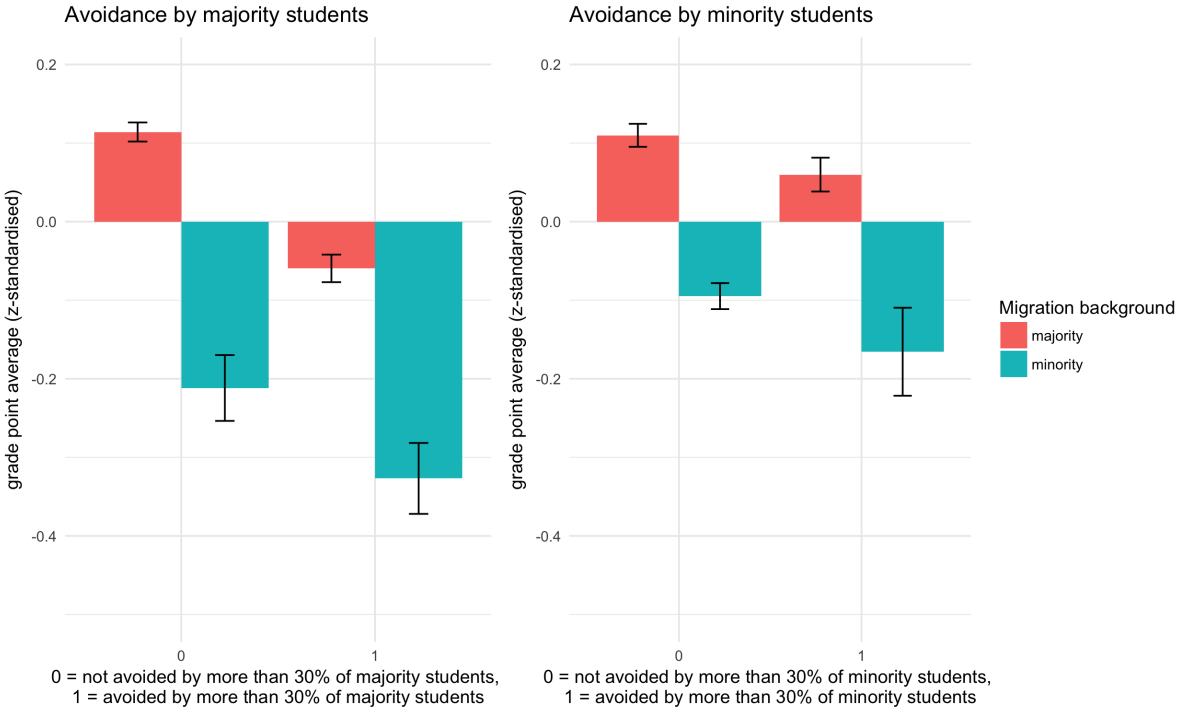

Figure 2.4: Grade point average by extent of social exclusion (avoidance) and migration background.

Table 2.2: Linear regression of grade point average (standardised).

	Model 1	Model 2	Model 3
migration background (ref: native)			
newly arrived	-0.55***	-0.36***	-0.30**
	(0.05)	(0.05)	(0.05)
1st generation	-0.22***	-0.14**	-0.12**
	(0.03)	(0.04)	(0.04)
2nd generation	-0.14***	-0.07*	-0.06
	(0.02)	(0.03)	(0.03)
gender (female)		0.09***	0.10***
		(0.02)	(0.02)
age (in months)		-0.01***	-0.01***
		(0.00)	(0.00)
cognitive ability		0.07***	0.07***
		(0.00)	(0.00)
highest parental degree (ref: primary school)			
lower secondary school		0.10**	0.08*
		(0.04)	(0.04)
upper secondary school		0.18***	0.17***
		(0.03)	(0.03)
university degree		0.54***	0.52***
		(0.03)	(0.03)
parental unemployment		-0.04*	-0.03
		(0.02)	(0.02)
other language spoken at home		0.07**	0.09**
		(0.03)	(0.03)
school resistance		-0.06***	-0.06***
		(0.00)	(0.00)
isolated from everyone			-0.09
			(0.16)
isolated from majority students			-0.12***
			(0.03)
avoidance by majority			-0.30***
			(0.07)
avoidance by minority			-0.03
			(0.04)
number of friends (indegree)			0.02***
			(0.00)
intercept	0.10***	1.05***	1.07***
	(0.02)	(0.26)	(0.26)
adj. R^2	0.01	0.21	0.22
N	10,275	10,275	10,275

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

All models control for survey country.

differences in effect sizes and significance levels (see Tables A.4 to A.6 in the Appendix): in England, the results are weakest, with non-significant but substantial effect sizes in the same direction, and in Germany strongest, with substantial and significant negative effects for social isolation from majority students and social avoidance by majority students. Sweden is in the middle range, with only avoidance by majority students having a negative effect.

Hypothesis 2 states that the negative effect of social exclusion accounts for immigrant disadvantages in educational success. This hypothesis is only partly supported. After including individual and parental controls (Model 2), the results show ethnic penalties for the first generation overall (though stronger for those who arrived at an older age), but not for the second, i.e. those born in the host country to immigrated parents. Where penalties are found, they decrease when controlling for social exclusion, but only marginally (and the reductions are not statistically significant). As the same general pattern shows for each survey country, the interpretation that social exclusion is behind some of the achievement differences is strengthened.

Hypothesis 3 states that the tendency of social exclusion accounting for immigrant disadvantages is strongest for those who migrated themselves. This is partly supported. In the model on the pooled data, social exclusion has most explanatory power when it comes to disadvantages of newly arrived immigrants. The country-specific models show the same tendency for England and Sweden, but in Germany, this also applies to second generation Turkish youths.

Hypothesis 4, stating that the negative effect of social exclusion on educational performance is worse for students with a migration background, is not supported (see Table 2.3[2]). The interaction effects show that, while including all controls mentioned as well as the other forms of exclusion, three measures of social exclusion have negative effect sizes, but none of the four is statistically significant.

When considering how these findings vary by survey country (see Tables A.7 to A.9 in the appendix), very similar results are observed: None of the interaction effects suggests that a particular exclusion measure is significantly worse for migrants than

[2]Results based on non-imputed data are shown in Table A.3 in the Appendix

Table 2.3: Linear regression of grade point average (standardised), with migration background and social exclusion interaction effects.

	Model 1	Model 2	Model 3	Model 4
migration background	-0.07*	-0.06	-0.09**	-0.07*
	(0.03)	(0.03)	(0.03)	(0.03)
gender (female)	0.15***	0.15***	0.14***	0.15***
	(0.02)	(0.02)	(0.02)	(0.02)
age (in months)	-0.01***	-0.01***	-0.01***	-0.01***
	(0.00)	(0.00)	(0.00)	(0.00)
cognitive ability	0.07***	0.07***	0.07***	0.07***
	(0.00)	(0.00)	(0.00)	(0.00)
lower secondary school	0.12***	0.12***	0.12***	0.12***
	(0.04)	(0.04)	(0.04)	(0.04)
upper secondary school	0.18***	0.18***	0.18***	0.18***
	(0.03)	(0.03)	(0.03)	(0.03)
highest parental degree (ref: primary school)				
university degree	0.54***	0.54***	0.54***	0.54***
	(0.03)	(0.03)	(0.03)	(0.03)
parental unemployment	-0.04	-0.04	-0.04	-0.04
	(0.02)	(0.02)	(0.02)	(0.02)
other language spoken at home	0.09**	0.09**	0.09**	0.09**
	(0.03)	(0.03)	(0.03)	(0.03)
isolated from everyone	-0.09	-0.09	-0.09	-0.09
	(0.10)	(0.09)	(0.09)	(0.09)
isolated from majority students	-0.11***	-0.00	-0.11***	-0.11***
	(0.04)	(0.11)	(0.04)	(0.04)
avoidance by majority	-0.38***	-0.38***	-0.53***	-0.38***
	(0.06)	(0.06)	(0.08)	(0.06)
avoidance by minority	-0.02	-0.02	-0.02	-0.02
	(0.02)	(0.02)	(0.02)	(0.02)
number of friends (indegree)	0.03***	0.03***	0.02***	0.03***
	(0.00)	(0.00)	(0.00)	(0.00)
isolated from everyone * migr.b.	0.01			
	(0.18)			
isolated from majority students * migr.b.		-0.12		
		(0.09)		
avoidance from majority * migr.b.			0.25	
			(0.13)	
avoidance from minority * migr.b.				0.01
				(0.10)
intercept	0.83*	0.83*	0.85*	0.83*
	(0.27)	(0.27)	(0.27)	(0.27)
adj. R^2	0.19	0.19	0.19	0.19
N	10,275	10,275	10,275	10,275

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

All models control for survey country.

Abbreviations: migr.b. = migration background.

it is for majority students. For Germany, however, the results suggest that the negative effect of majority students avoiding the individual is weaker for minority students compared to majority students.

2.4.3 Robustness checks

A critique for the measure of social isolation is that an individual might have all friends outside the classroom. In addition to the full friendship network for each classroom, the *CILS4EU* data includes information on each individual's five best friends from an ego-network perspective. This means that friends that are nominated in this way are not identified, and not in turn asked to nominate friends as well. That, however, also implies that anyone can be named, whether in the school class or not. Each respondent is then asked to give additional information about each friend, such as their gender, their ethnicity, and whether they are in the same school and class or not. Based on this, the variable for social isolation was redefined in such a way that it only applies to individuals who are socially isolated in the classroom and who do not report friends outside school. Likewise, the variable for social isolation from majority students was redefined in such a way that it only applies to those who have no majority friends inside and outside school. Analyses with both versions of the isolation variable were carried out, and the substantial conclusions remain the same: It does not make a difference if individuals have friends outside school if they are isolated in the classroom. Conceptually, this is a sensible finding, since friends outside the school are not able to help with homework or exam preparations, and might generally be in a different life situation.

As discussed above, social avoidance ties were only counted as such, if the individual did not consider the sender to be their friend. Since there are various reasons for not wanting to sit by someone that they otherwise consider their friend, for example not being able to concentrate on the lessons, analyses were carried out with both versions of the variable. The substantial results are stronger when social avoidance ties coming from friends are excluded. This implies that avoidance ties sent by friends are actually tapping into a different dimension than those from non-friends when it comes to their relationship to grade point average: It is likely that friends, even if they do not want to

sit next to the individual, can still be valuable resources of support and advice. While this certainly can be different in other cases, these theoretical arguments strengthen the decision on how to deal with this in the context of this study.

2.5 Discussion and conclusion

The main finding of this study is that social exclusion is related to lower grade point average, which applies to both majority youths and those with a migration background. Although various concepts of social exclusion were considered, and a tendency for a negative association with GPA was seen throughout the analysis, the only significant results were found for those measures that capture exclusion from the majority students in the classroom: being isolated from majority youths, and being avoided by majority youths. This implies that access to "native social resources" matters more, which is in line with theoretical considerations on the beneficial knowledge on manoeuvring the educational system that majority youths are more likely to possess, relative to those with a migration background, as well as to the value of classmates' help with the school work. The possibility of reverse causality cannot be excluded here, so that the association reflects the avoidance of class-mates who are underachievers, possibly because of their behaviour, but the analysis manages to control for a number of items indicating school resistance, such as skipping classes and arguing with teachers. The analysis did not show, however, that being socially excluded is worse for migrants, although this test had to be crude to maintain statistical power.

This study set out by asking whether social exclusion can account for some of the disadvantages in grades that children of immigrant origin experience in relation to their peers of majority background. Although we did find some signs of this, the impression is that this hypothesis was not supported. While exclusion is important for grades, and more children of immigrant origin are socially excluded, the relations are not strong enough to explain the grade difference that was found.

There are limitations to this research. While it is assumed that the mechanism that is investigated here is a general one, the analysis could only include data from three countries, and all of them are European. Ethnic composition of the migrant population

varies due to different migration histories of the countries. Since social network data is not a standard feature of large-scale cross-national surveys, future analyses that consider other national contexts depend on the collection of more data of this kind.

Furthermore, this analysis can only point to correlations, while the process linking social exclusion to achievement is likely to start before youths entered secondary school. Collecting and analysing longitudinal data on significant peers and the family from an early age would rectify this limitation. Lastly, more research is needed on the actual mechanism on how social exclusion works, to support and help interpret the initial results that are found in this paper. Next to limited access to social resources, lower mental-wellbeing is another potential pathway by which social exclusion could negatively affect school grades.

However, there are various reasons that render the framing of this study plausible and its findings valuable. While the cross-sectional analysis limits the claims that can be made, there are strong theoretical foundations that build on various strands of previous research, as presented in the literature review. The social dimension is important to include in considerations of educational success, and students who are socially excluded are likely to lack valuable resources, which plays out negatively in their educational success. Most of these combined theoretical implications have not been tested explicitly in the past, due to unavailability of suitable data. Using a large-scale cross-national survey that includes social network data, and can thus accurately measure social dimension of educational success, renders this study particularly valuable.

Chapter 3

Social milieus and educational aspirations

This study examines how educational aspirations among adolescents are clustered within school classes, and whether this differs for students with a migration background. Educational aspirations have long been found to be of relevance when explaining educational success and diverging patterns of ethnic penalties. Through the utilisation of social network measures from the *CILS4EU* data on Germany, the Netherlands, and Sweden (15,203 individuals, 50.04% girls, and on average 14.9 years old), we are able to measure the peer dimension more accurately in comparison to studies that merely use classroom-level aggregates. Using longitudinal logistic regression models and first-difference models with individual-level fixed effects, we find that adolescents cluster within classrooms based on educational aspirations, and that this pattern is less pronounced for individuals with a migration background.

This chapter is based on the paper "What is going on around you: Social milieus and educational aspirations", that is currently revised for publication at *European Sociological Review*. It is co-authored with Ralf Wölfer, Department of Experimental Psychology, University of Oxford.

3.1 Introduction

In research on differences in educational success, various factors have been proposed to be of relevance, with the traditional predictor being social and economic resources (Erikson and Goldthorpe, 1992). One of the areas in which these can play out are educational aspirations, i.e. the educational and occupational career the individual aspires to, which can affect educational outcomes, for example through higher motivation or goal orientation (Sewell et al., 1969; Erikson and Jonsson, 1996; Breen and Jonsson, 2005).

Aspirations do not manifest themselves out of nowhere, but are socialised entities. This is a dynamic process that begins within the family, which is the reason why the socioeconomic or migration background of the family is often found to have an effect on aspirations. Throughout adolescence, socialisation increasingly happens through the additional contact with peers. In fact, it is assumed that for adolescents, peers (i.e., others of the same age) provide the more significant frame of reference compared to parents, due to time spent together and similar life circumstances (e.g. Youniss and Smollar, 1985; Osterman, 2000). Peer effects have thus been extensively researched in the past, in a variety of areas of the social sciences.

Since the publication of the "Coleman Report" on equality of educational opportunity (Coleman et al., 1966), educational outcomes have been a focus of peer research, too. While this strand of research has brought forward insights on attainment and performance, we propose to consider educational aspirations more specifically, which has been comparatively neglected in research in the past. Since aspirations are a mental orientation towards a specific goal, we propose that this is more easily changeable and adjustable than more finite outcomes, such as the degree attained or the grade point average achieved. Furthermore, as illustrated above, aspirations are one of the mechanisms that probably influence educational outcomes.

Our research aims at showing that the peer dimension matters for educational aspirations. Given the extensive research showing the strong relationship between aspirations and educational success, understanding the dynamics of aspirations is an important factor in explaining diverging patterns of educational success. This is especially

true for students with a migration background, since it has been found that parental and individual aspirations among some ethnic groups are comparatively higher, which has strong effects on educational outcomes (Heath et al., 2008; Jonsson and Rudolphi, 2011). Thus, our findings also have important policy-implications.

We furthermore add to existing research by deploying a more explicit way of measuring the peer dimension, in comparison to traditional sociological and educational studies. Although these studies assume mechanisms that involve dynamic interactions with relevant others, this theoretical assumption is usually not tested explicitly. Instead of measuring what is happening in the immediate social surrounding of the individual, researchers often use contextual factors, which are aggregated on the level of the classroom or the neighbourhood (e.g. Legewie and DiPrete, 2012, 2014). In this paper, we advance the recent literature by measuring the immediate social surrounding through social network data, to get at the exact dimension that socialisation theory assumes to be of relevance when it comes to the dynamics of aspirations. By carefully applying a regression framework to these measures, we are able to draw novel conclusions on peer clustering and educational aspirations from large-scale cross-national data. We thus contribute to existing literature by bridging the gap between theory and empirical research on peer effects in mainstream sociological research. We are aiming at answering the following research questions in this study: To what extent do educational aspirations cluster within classrooms, i.e. on the peer level? Does this differ for students with a migration background?

3.2 Theory, previous research, and hypotheses

This section gives an overview of relevant theory and previous research, and develops the hypotheses guiding the research presented in this chapter.

3.2.1 Why do aspirations matter for educational success?

Educational success, whether it is measured through grades or level of highest educational degree, has been shown to be impacted by both individual as well as contextual circumstances. For example, on the individual level, the probability of achieving a

degree from a tertiary education institution is highly correlated with the economic resources of the family in general, as well as parents having a university degree. Reasons for this are both social and human capital, i.e. resourceful connections to others or strategic knowledge about the education system (Corak, 2013) that spread through the network. One example of how the contextual level matters, is the so-called big-fish-little-pond effect (e.g. Marsh et al., 2001), indicating that the contextual reference point measured by the average skill or achievement level of students' classmates negatively influences students' academic self-concept.

Another way in which the individual and contextual level influence educational success is through the educational aspiration that students hold for themselves. In particular, this refers to the specific educational and occupational career path the individual aims to follow, and can thus be connected to the ambition of attaining a particular educational level, or of achieving particular grades. Various studies in different branches of the social sciences have documented a strong effect of aspirations on educational and occupational outcomes (Feliciano and Rumbaut, 2005; Portes et al., 2010). This is mainly explained by the potential of aspirations to increase motivation and the tendency to make more ambitious career choices. Aspirations are thus affecting educational performance and choices which are considered to explain educational outcomes. These two aspects, which are also known as primary and secondary effects on educational attainment, have received much attention in social research in the past decades (e.g. Erikson and Jonsson, 1996; Jackson, 2013). Since educational success carries implications for level and type of occupation the individual will be eligible for, aspirations continue their significance for labour market outcomes as well.

Aspirations have also been found to explain the tendency of adolescents with a migration background being more successful in school than their native peers from the same socioeconomic background that has been documented in some national contexts (Heath et al., 2008). This is generally explained by migrant parents having high aspirations for their children, because they view the migration event to be part of a social mobilisation project of the entire family (Heath et al., 2008) or because they originate from a social stratum in their previous home country, which is comparatively higher

than the one in the destination country (van Tubergen and van de Werfhorst, 2007; Ichou, 2014). When these higher aspirations are transmitted to their children, we see groups of adolescents with a migration background who on average have higher educational aspirations than their native peers, even when they perform equally well in school (Brinbaum and Cebolla-Boado, 2007; Jackson et al., 2012).

Various factors have been found to affect educational aspirations, which are related to family background, such as socioeconomic as well as migration background. There is, for example, the idea that values regarding the worth and extent of education differ between social classes. In particular, this refers to children from upper or middle classes being socialised into a value system that puts higher regard on education, in comparison to children who grow up in a working class. This can play out in different risk-avoiding behaviours regarding the career, such as taking out student loans, or following a career path which does not lead to a pre-defined occupation, but has higher earning potential (Jonsson and Rudolphi, 2011).

3.2.2 Peers and aspirations

Findings from a variety of branches of social sciences have identified peers to be of importance when it comes to individual outcomes during adolescence. Examples exist for both negative and positive outcomes, such as substance abuse (for reviews see Borsari and Carey, 2001; Kobus, 2003), deviant behaviour (e.g. Dishion and Tipsord, 2011), depressive symptoms (Van Zalk et al., 2010), and eating disorders (Rayner et al., 2013) on the one side, and prosocial behaviour (Barry and Wentzel, 2006; Wentzel et al., 2004), higher emotional well-being (Elmer et al., 2017), and empathy (Wölfer et al., 2012) on the other side.

Educational outcomes have been assumed to be influenced by peers for several decades. As mentioned above, the influential "Coleman Report" (Coleman et al., 1966) was one of the first pieces of research proposing that individual success might be subject to the average success of others in the school class. Since this early research, various studies have been put forward. Although there is some variation in the strength of the results, the overall tendency of this research is that peers matter for educational

outcomes. It has been found, for example, that students who study together with their peers perform better in school; the same applies to more abstract forms of support, such as emotional and social support (Wentzel et al., 2010; van Rijsewijk et al., 2016). Furthermore, it has been found that the orientation of others towards education as such can also have an effect on the individual, if a normative peer culture exists in the learning context. For example, a general sense of fairness and control over one's educational success in a classroom is related to individual achievement (Agirdag et al., 2012). A similar concept, known as the "acting white" phenomenon, has also been researched extensively in the United States. It refers to a learning culture, that can be more or less present in a learning context, in which Black students consider educational success to be a white privilege, which therefore should not be aimed at. The prevalence of this attitude in a learning context has been found to be detrimental to individual educational outcomes (Fordham and Ogbu, 1986; Fryer and Torelli, 2010).

In this study, we are zooming in on one mechanism that links individual characteristics to educational outcomes: educational aspirations. While the traditional literature on status attainment mainly focused on the effects of family background on individual outcomes (Blau and Duncan, 1967), the so-called Wisconsin model of status attainment explicitly began to consider aspirations, too (Sewell et al., 1969). Next to individual aspirations, they emphasise the positive effect of "significant others", such as parents' and teachers' encouragement and friends' own aspirations on educational and occupational attainment. In a field experiment on policy-induced social interactions and schooling decisions in Mexico, researchers found evidence of how social contact can affect educational decisions, and their findings indicate that this is due to increased educational aspirations (Bobba and Gignoux, 2011). Furthermore, there are some studies documenting the general spread of norms or attitudes through the social network (e.g. Friedkin, 2001). Considering these previous findings on peer effects on educational outcomes and the fact that aspirations are assumed to be subject to similar dynamics, we expect peers also to be of importance for educational aspirations.

As mentioned above, educational aspirations are assumed to be one of the main drivers of the observed decoupling of socio-economic background and educational out-

comes for migrants, which is why we are interested in whether peer effects on aspirations differ by migration background. As argued above, ethnic variation in educational success has been explained by differences in aspirations. This is especially the case for some groups of second generation migrants in Europe that outperform the native majority; these differences are assumed to be largely due to the higher motivation of those whose family has migrated. These higher levels of motivation have mostly been considered as being part of the family background, but, as argued above, those of the same age become more important in the course of adolescence. Therefore, we ask whether they are as important for youths with a migration background as they are for natives. This procedure opens many possible paths of exploration in our area of research, such as considering intra- and interethnic friendships separately or including measures of segregation in the classroom, investigating these alleys more closely is beyond the scope of this paper though. We are thus focusing on exploring if and how peer dynamics differ between natives and youths with a migration background, to build a platform and to provide guidance for fruitful future research in the area of integration and segregation.

3.2.3 Friends instead of classmates

While peer effects have a long tradition in social research, especially in psychology and economics, the usual analytical strategy does barely consider peer effects as such. Usually, we see the peer effect conceptualised as an aggregate of characteristics of a wider social context, such as the school class of the individual, or the neighbourhood (e.g. Rjosk et al., 2014; van Ewijk and Sleegers, 2010). We argue that this operationalisation is improvable. Individuals have a need for social connectedness and close friendships, which they cannot uphold with everyone in their social context to the same extent (Baumeister and Leary, 1995). There are comparatively few people in these settings that we spend the majority of our time with, and they are likely to be most relevant to how we adjust and develop our attitudes and behaviour. It is therefore more appropriate to focus on this group in the social context when researching peer effects. One way to implement this is to focus on students' explicit friends or the meso-level that exists between the individual and the larger context, which is our approach in this paper.

Research on friends has repeatedly shown the tendency of homophily, i.e. people becoming friends with those that are similar to them as to a vast variety of characteristics (McPherson et al., 2001). There are several theories that aim at explaining this often-observed pattern, such as Similarity Attraction Theory (Byrne, 1997; Kossinets and Watts, 2009), which proposes that it is more rewarding to be friends with someone who is similar to oneself. This higher reward is attributed to easier communication and improved understanding due to similarity, which then leads to higher and more easily achieved levels of trust and solidarity. Homophily is likely to exist regarding educational aspirations, too. Sharing the same educational aspiration, which might play out in taking school lessons seriously or exhibiting goal-oriented behaviour, can facilitate communication and empathy towards one another, which is an important basis for friendship. Furthermore, this can lead to shared experiences, from initially preparing together for an exam and attending a career event, to later spending also free time together.

Simultaneously, being friends can also lead to an alignment of educational aspirations, for example through the mere exposure to the idea of going to university. Having a friend who is in a similar situation to oneself and who has a specific aspiration can turn an abstract plan into a definite one, since it seems more realisable, or because specific pieces of information are conveyed. Both tendencies, alignment of aspirations due to friendship, and friendship due to initial homophily, result in clustering of educational aspirations in the friendship network of a classroom. By aggregating educational aspirations on the class-level, this variation between individuals of a class is not accounted for.

3.2.4 Social milieus in classrooms

In this study, we are using friendship network data to look at the more immediate social surrounding of the individual, which we are conceptualising as the "social milieu". We regard the social milieu to be nested within the wider social context, i.e. the school classroom in our instance. In doing so, we are able to account for the exact mechanism that socialisation theory assumes to be of relevance when it comes to educational

aspirations, in contrast to the vast number of studies that use more coarsely grained measures of the peer context. The first hypothesis that we are testing in this research is thus:

Hypothesis 1: Educational aspirations are clustered in social milieus.

The social milieu can be conceptualised in two ways: as friends and as friendship groups. Friends are simply those that individuals consider to be their friends. Friendship groups, as they are understood in this context, are informal groups within classrooms, whose members are frequently exposed to each other, but who do not necessarily all consider each other to be friends, since we include one-sided friendships as well. In other words, friendship group members are tied to each other by at least one individual considering the other one to be a friend. There are several studies that have identified a tendency for the existence of unbalanced friendships in friendship groups. This has been traditionally interpreted as a sign of hierarchy between two of the individuals (Davis, 1970), but more recently has been explained by the tendency that friendships embedded in social groups are more stable, and can thus also exist in the long-term if they are one-sided (Block, 2015).

In differentiating between these two concepts, direct friends and friendship groups, we are measuring slightly different dimensions of the social milieu of an individual. While those that are regarded as friends are likely more influential when it comes to adjusting aspirations based on their own ones, and they are likely more similar to the individual to begin with, as suggested by the tendency of homophily. Friendship group members can also be friends, but they do not necessarily have to be. Thus, an individual can be less able to choose whether they want to spend time with each member of the group, but they are nevertheless likely to be frequently exposed to each other. Additionally, the concept of the friendship group entails a situation, in which specific norms and attitudes are likely to be negotiated and renegotiated more frequently. We thus expect both dimensions of the social milieu to be both equally relevant when it comes to clustering of educational aspirations.

Hypothesis 2: The clustering of educational aspirations will be found across different operationalisations of the social milieus, i.e. on the level of friends and on the level of friendship groups.

It can be expected that classrooms are relatively homogeneous in terms of educational aspirations. Additionally, the classroom is also a part of a wider social system, such as the school. It can be assumed that individuals on average interact more with others in their class than with others that are not in the same class but in the same school. Thus, we are furthermore expecting clustering based on the classroom. This between-classroom variation, however, cannot completely account for the within-classroom variation, or in other words the clustering of aspirations within classrooms.

Hypothesis 3: Social milieu remains important even when we consider the broader contextual effects.

As discussed above, previous research suggests that higher educational aspirations of migrants can largely be explained by family characteristics, which are related to the migration event, such as aiming at increased social mobility in the new home country. Since we see this playing out in the equalisation of educational success for some migrant groups, it is likely that youths with immigrant background can be very much affected by these family mobilisation aspirations. We therefore expect peers to play a lesser role for migrants, when it comes to aspirations.

Hypothesis 4: Clustering of aspirations is less pronounced for adolescents with a migration background.

The main strength of our analysis is our ability to measure directly the theoretically relevant dimension of the social milieu by the utilisation of social network data. We are able to show correlation and clustering of educational aspirations of individuals within

a large number of classrooms in three European countries, and can link our findings to previous literature on peer effects in education. We furthermore aim at illustrating the value of a longitudinal social network analysis in this area, by exploiting the potential of the analytical framework at hand, and by highlighting its limitations.

3.3 Data and Methods

This section introduces the data and methods used in the analysis.

3.3.1 Data

This paper uses waves 1 and 2 from the Children on Immigrants Longitudinal Survey in Four European Countries (*CILS4EU*) dataset (Kalter et al., 2014, 2015). It is fully standardised, internationally harmonised, and consists of surveys completed by the students at age 14-15 (wave 1) and 15-16 (wave 2), their parents, and teachers. Furthermore, it contains socio-metric, i.e. network data both on positive ties (friendships) and negative (rejection) ties.

This paper uses the survey data on the individual student, as well as the data on the friendship networks in the classroom, to construct measures for the social milieu, as described below. Since the socio-metric part of the data from England has to be treated with caution due to some errors in the data collection, we are focusing on the Dutch, German, and Swedish part of the data. We excluded classes with less than 15 students in order to analyse reasonably large networks.

3.3.2 Variables

Table 3.1 shows descriptive statistics of all variables in the full sample and the analytical sample. There are no substantial differences in the descriptive statistics of the two samples, although we lose a number of cases. We are unable to use multiple imputation to deal with item non-response, since our main variables of interest are based on social network data. Imputation of network data is not straightforward and is currently only under development. We are thus losing a relatively large number of cases. To investigate whether we have systematic attrition in our sample, we predicted being in the analytical

sample by the individual level characteristics which are also included in our analysis; the results can be found in Table B.1 in the Appendix. We see several dimensions that are significantly predicting presence in the analytical sample, yet, the effect sizes are small. Since the number of cases is very high in the model, resulting significance levels in this test could stem from this. Overall, the low effect size together with the descriptive results gives us reason to assume that we do not see systematic attrition in our sample that could bias the results. We used multiple imputation for variables that are not based on social network data (10 imputed data-sets). Results based on non-imputed data can be found in Tables B.2 and B.3 in the Appendix. Analyses based on imputed data and list-wise deletion yield the same substantial conclusions, with only little variations in effect sizes and standard errors.

Outcome variable: Academic aspirations of the individual

The outcome variable, educational aspirations, is based on the survey item "What is the highest educational qualification you wish to get?". Four answers were possible: no degree, degree below upper secondary school, degree from upper secondary school, and university degree. In Sweden and the Netherlands, there were only a few that reported that aspirations in the lowest two categories (in Sweden less than 1%, and in the Netherlands 8%, in wave 1), so the lower three categories were merged. In the German part of the data, there was more variation (less than one percent named the lowest category, 27% named the second lowest, and 39% the second highest). We ran the analyses with both a binary and an ordered outcome, both returning the same substantial conclusions. To facilitate the interpretation of our overall analysis, we thus also merged the lower three categories for the German part of the data. Overall, students in Germany report the lowest educational aspirations, and in Sweden the highest aspirations, and students overall tend to report higher aspirations at wave 2.

Predictor variables: Academic aspirations in the social milieu

As mentioned above, we measure the social milieu in two ways, as friends and as friendship groups. For each individual, we compute the social milieu, i.e. their friends

Table 3.1: Variable and sample descriptives: full and analytical sample.

	Full sample						*Analytical sample*					
	Mean	SD	Median	Min	Max	N	Mean	SD	Median	Min	Max	N
Individual university aspirations, w1	0.51	0.50	1.00	0.00	1.00	14314	0.55	0.50	1.00	0.00	1.00	7147
Individual university aspirations, w2	0.58	0.49	1.00	0.00	1.00	12374	0.62	0.48	1.00	0.00	1.00	7147
Age (in months)	179.91	7.59	179.00	157.00	228.00	14342	178.76	6.75	178.00	158.00	219.00	7147
Cognitive ability index	18.57	4.46	19.00	0.00	27.00	13984	19.16	4.16	20.00	0.00	27.00	7147
At least one parent has a university degree	0.25	0.43	0.00	0.00	1.00	15203	0.29	0.45	0.00	0.00	1.00	7147
Parental university aspirations	0.38	0.48	0.00	0.00	1.00	14300	0.41	0.49	0.00	0.00	1.00	7147
Friends: av. university aspirations, w1	0.51	0.36	0.50	0.00	1.00	13417	0.55	0.34	0.60	0.00	1.00	7147
Friends: av. university aspirations, w2	0.62	0.36	0.67	0.00	1.00	9874	0.63	0.36	0.67	0.00	1.00	7147
Friends: av. cognitive ability	18.69	3.18	19.00	0.00	27.00	13375	19.02	3.00	19.40	1.00	26.50	7147
Friends: prop. with parental degree	0.25	0.28	0.20	0.00	1.00	13460	0.28	0.29	0.20	0.00	1.00	7147
Friends: prop. natives	0.63	0.36	0.75	0.00	1.00	13424	0.66	0.35	0.75	0.00	1.00	7147
Friendship group: av. university aspirations	0.50	0.35	0.40	0.00	1.00	13862	0.54	0.38	0.50	0.00	1.00	7147
Friendship group: prop. with parental degree	0.25	0.30	0.17	0.00	1.00	13878	0.29	0.33	0.20	0.00	1.00	7147
Friendship group: av. cognitive ability	18.50	3.10	18.00	0.00	27.00	13835	19.00	3.32	19.33	1.00	27.00	7147
Friendship group: prop. natives	0.62	0.35	0.56	0.00	1.00	13878	0.66	0.38	0.75	0.00	1.00	7147
Class: av. university aspirations (- friends), w1	0.38	0.21	0.39	0.00	0.94	14314	0.40	0.19	0.41	0.00	0.93	7147
Class: av. university aspirations (- friends), w2	0.39	0.20	0.41	-0.20	0.95	12374	0.41	0.18	0.42	-0.06	0.95	7147
Class: prop. with parental degree (- friends)	0.20	0.16	0.17	0.00	0.83	15203	0.21	0.15	0.18	0.00	0.81	7147
Class: av. cognitive ability (- friends)	13.44	2.97	13.58	0.00	22.56	13984	13.65	2.63	13.67	5.44	21.58	7147
Class: prop. native students (- friends)	0.69	0.22	0.75	0.04	1.20	14420	0.74	0.20	0.79	0.07	1.06	7147
Class: av. university aspirations (- friends)	0.53	0.26	0.56	0.00	1.00	13390	0.54	0.25	0.57	0.00	1.00	7147
Class: av. cognitive ability (- fr.group)	18.71	2.31	18.83	8.17	23.85	13367	18.85	2.30	18.94	8.17	23.85	7147
Class: prop. with parental degree (- fr.group)	0.26	0.19	0.22	0.00	0.94	13390	0.26	0.19	0.24	0.00	0.94	7147
Class: prop. native students (- fr.group)	0.64	0.26	0.71	0.00	1.00	13390	0.65	0.26	0.73	0.00	1.00	7147
Country: Germany						5127						2235
Country: The Netherlands						4628						2101
Country: Sweden						5448						2811
Gender: male						7192						3463
Gender: female						7204						3684
Migration background: native						8991						4729
Migration background: non-native						5429						2418

and those in their friendship group, and take the proportion of those who aspire to a university degree. The measure of friends is based on the network survey item "Who are your best friends in the school class", for which five nominations were possible, and which exists both at wave 1 and 2. Thus, it is a measure strictly from the perspective of the individual, which does not account for relations between their friends.

Our second measure for capturing the social milieu represents the friendship groups in the social network of students (for a graphical illustration of a sample network in the used dataset, see Figure 3.1). Friendship groups were extracted by applying the Hierarchical Clique Clustering Approach (Everett and Borgatti, 1998) for wave 1 of the *CILS4EU* data. A clique is methodologically defined as a group of people within a larger network, between whom all possible ties exist (Luce and Perry, 1949). For our case, this would be a group of people in a classroom, who all nominate each other as friends. This, however, usually results in overlapping groups within classrooms, which is not desirable for our purposes: we want to measure the one most important social context for each individual. The HCCA is a tool for extracting this most important friendship group for each individual.

It starts out by computing cliques that overlap, i.e. it determines sub-networks of students, within the full classroom network, who all nominated each other as friends. Following our theoretical considerations on groups and unbalanced friendships, we apply the HCCA to the undirected friendship networks in the classrooms, thus also considering one-sided friendship nominations as a connection between those two individuals. It is then analysed how these cliques cluster, in other words, the network of cliques is then investigated further. In practice, this is done through formulating a co-clique matrix, in which the entry of a cell (i, j) is the number of students that clique i and clique j have in common. The standard hierarchical clustering routine in *UCINET* then assigns each clique to a unique cluster. This means that clusters are merged, reducing but not necessarily eliminating overlap between cliques. We consider the cliques resulting from this analysis to be the most important friendship group that an individual is embedded in. In classrooms for which a non-overlap could not be achieved, we were unable to compute the friendship group dimension of the social milieu, and thus excluded those

classrooms from the analysis.

The measure of academic aspirations in the social milieu is based on the proportion of people in the social milieu who have university aspirations, leaving out the individual. Thus, for each student in our data, there exists a measure of the academic aspirations of the other students for both dimensions of the social milieu: one based on friends, one based on friendship groups.

To understand the extent to which the two theoretical concepts of the social milieu actually differ in practice, the overlap of classmates that were captured by the two different measures was calculated. Each individual considers on average 11.8% friendship group members not to be his or her friend. Considering that we include one-sided friendship nominations, this means that group members who were not nominated as friends have nominated the individual in question. Given that only five friendship nominations were possible, this could also be due to the fact that respondents were simply not able to name all of their friendship group members as friends. Thus, we also consider this overlap only for those who nominated less than five friends, implying that they would have had the space to nominate the other group members as well. With 12.7%, the average percentage of non-friends in the friendship group is slightly higher for this sub-sample. Analogously, we calculate the proportion of nominated friends who are not in the respective friendship group. On average, this applies to 13.9% of an individual's friends. These proportions are very similar, and they are in line with the theoretical considerations mentioned above, i.e. that the two measures of the social measures are slightly different, yet likely similar in their conceptual importance for aspirations.

Predictor variables: Academic aspirations in the classroom and migration background

Based on research that evidently shows that the larger context besides peer groups has an effect as well, we are also including academic aspirations in the classroom in our research strategy. To account for the dynamics on the classroom level, net of the dynamics in the social milieu, we leave out the social milieu when computing the

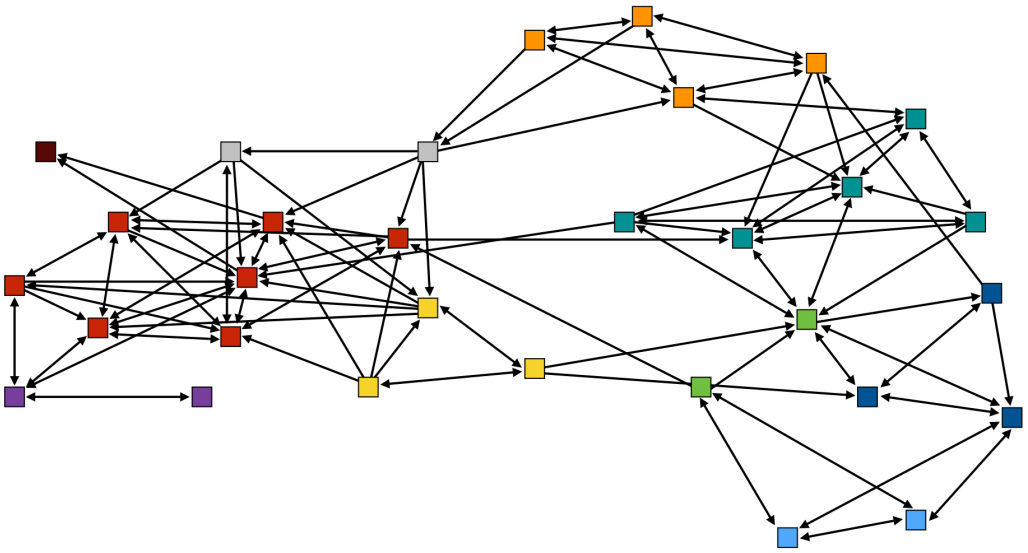

Figure 3.1: Identified friendship groups based on Hierarchical Clique Clustering Approach (HCCA) in a sample classroom.

proportion of students in the classroom who aspire to attaining a university degree. It follows that we have two measures of the classroom context for each individual, one aggregated measure on the classroom level net of the friends and one net of the friendship group of an individual. The individual's own aspirations are also not included in the class-level contextual measure of university aspirations, either.

Migration background is included in the analysis as a dummy variable. We consider those individuals who migrated themselves or who have at least one parent who migrated to have a migration background.

Controls

We include a number of controls in our models. The proportion of migrants in the school is included to control for the sampling stratum, and is necessitated by the survey design, where immigrant-dense schools were over-sampled. We furthermore control for the survey country.

On the individual level, we control for a variety of confounders in relation to academic aspirations. As discussed above, socio-economic background is usually a strong predictor of educational aspirations. Since we are concerned with academic aspira-

tions, we are including a binary measure of whether at least one of the parents has a university degree. In line with the above-discussed Wisconsin Model, we control for parental aspirations, i.e. whether parents want their child to get a university degree or not. We furthermore control for cognitive ability of the individual, to account for the possibility that ability affects both aspirations and motivation, which could introduce a spurious effect of aspiration. This control is based on a test implemented as part of the *CILS4EU*. It is language-free, since it is based on solving graphical (pattern) problems in a given time (CILS4EU, 2014). Since cognitive ability depends to a large extent on parental background (e.g. Heckman, 2006), this is also likely to pick up further effects related to socio-economic background. We moreover control for gender as well as age in months at the time of the survey.

On the level of the social milieu, we control for the proportion of others with at least one parent that has a university degree, the proportion of native others, as well as the cognitive ability of others. We focus on natives among the explicit friends as well as in the friendship group instead of those with a migration background, since this category includes a variety of different ethnic groups, and can thus refer to varying compositions in different instances. In line with how we compute educational aspirations in the social milieu, we also calculate this as the mean, leaving out the individual.

Correspondingly, we control for the same factors on the contextual level, while also leaving out the individual and the social milieu in the computation. By including proportion of natives, average cognitive ability, and proportion of students with a parental university degree of others on the contextual level, we aim to account for sorting into classrooms which will likely explain clustering of educational aspirations on the class level as well.

3.3.3 Methods

We use two main types of analyses in our approach. First, we consider a first-difference model using individual-level fixed effects. Here, we are predicting the change in individual aspirations by the change of aspirations in the social milieu, as well as the change in school class aspirations, between waves 1 and 2, to analyse whether and how

individuals' aspirations change with the corresponding change in their milieu. We use only the change of aspirations in the social milieu of explicit friends, since we cannot account for endogenous network processes such as reciprocity and transitivity, which affect the change of the overall network structures between wave 1 and 2. HCCA, the technique to compute our social measure of friendship groups, uses exactly these pieces of information, which is why we focus on friends. The specification of a first-difference model with individual-level fixed effects only allows for within-individual comparisons, which eliminates omitted variable bias, i.e. unobservable traits across individuals that affect both the predictor and the outcome variables. This is achieved by eliminating all time-invariant observed and unobserved characteristics, i.e. those that do not change across surveys (Allison, 2009). Model 1 of the first-difference model includes only the change of aspirations in the social milieu as an independent variable, and model 2 adds the change of aspirations in the classroom to model 1. We are interested in how migration background applies to this dynamic of change, but since the migration background is a time-invariant individual trait and thus does not change between waves 1 and 2, we cannot include this in the model as an independent variable. However, in order to study whether the social influence on aspirations differs by migrant status, we also calculate the model separately for migrants and non-migrants.

While the first-difference model with individual level fixed effects has a lot of power in analysing change, it cannot account for between-group variance in the overall data and thus cannot explain initial sorting and clustering. We thus include two sets of logistic regression models, predicting university aspirations at wave 2 by the social milieu at wave 1: one set of models with the social milieu of friends, and one set of models with the social milieu of friendship groups. Both sets are built up in the same way. We start with a simple model which includes only survey country and sampling stratum besides the main independent variable of interest: the university aspirations of others in the social milieu. Model 2 then adds all individual and social milieu-level controls, and Model 3 the contextual-level predictors and controls. Model 4 adds a migration background interaction effect. We report average marginal effects (Mood, 2010).

All analyses have been carried out in *R* 3.3.2, using the *plm* package for the regression analyses.

3.4 Results

Before we present our results from the various regression models described above, we briefly discuss some descriptive results on the main variables of interest: educational aspirations, the social milieu, and the migration background.

3.4.1 Descriptive results

When considering the change in aspirations by migration background, we see no differences between natives and ethnic students. Of all native students, 18% report different aspirations in wave 2 in comparison to wave 1. For ethnic students, 17% change their aspirations between wave 1 and 2. The direction of change is also very similar: 64% of all change of ethnic students is upwards; for natives, it is 65% of all change.

Although we are not explicitly focusing on the difference between inter- and intra-ethnic friendships in our analysis, we do consider the proportion of natives among the best friends and in the friendship group: Overall, we see that migrants have substantially fewer natives in their social milieu. For the proportion of native friends among the best friends, there is not much difference between the two waves. In wave 1, natives have on average 76%, and in wave 2 77% natives among their best friends. Ethnic students have on average 41% natives among their best friends, in both waves. The proportion of natives among others in the friendship group (measured only at wave 1) is slightly different, but still points to the same tendency. For natives, 73% of others in the friendship group are also natives, whereas ethnics have on average 44% natives in their friendship group.

3.4.2 Results from regression analyses

Tables 3.2 and 3.3 show the results from the logistic regression models predicting individual university aspirations at wave 2, considering the friends and the friendship

Table 3.2: Logistic regression predicting individual university aspirations at wave 2 (reporting average marginal effects, standard errors clustered on class level). Friends as the social milieu.

	Model 1	Model 2	Model 3	Model 4
friends: av. univ. aspirations	0.49***	0.18***	0.12***	0.12***
	(0.02)	(0.02)	(0.02)	(0.02)
individual univ. aspirations, w1		0.39***	0.38***	0.38***
		(0.01)	(0.01)	(0.01)
gender (female)		0.06***	0.06***	0.06***
		(0.01)	(0.01)	(0.01)
age (in months)		-0.01***	-0.00***	-0.00***
		(0.00)	(0.00)	(0.00)
migration background (vs. native)		0.09***	0.09***	0.09***
		(0.02)	(0.02)	(0.03)
cognitive ability index		0.01***	0.01***	0.01***
		(0.00)	(0.00)	(0.00)
at least one parent has a univ. degree		0.07***	0.06***	0.06***
		(0.02)	(0.02)	(0.02)
parental univ. aspirations		0.16***	0.16***	0.16***
		(0.02)	(0.02)	(0.02)
social milieu: friends				
av. cognitive ability		-0.05	-0.05	-0.05
		(0.03)	(0.03)	(0.03)
prop. natives		0.01***	0.01**	0.01**
		(0.00)	(0.00)	(0.00)
prop. parental degree		0.14***	0.13***	0.13***
		(0.03)	(0.03)	(0.03)
school class				
av. univ. aspirations			0.35***	0.35***
			(0.06)	(0.06)
prop. natives			0.09	0.09
			(0.06)	(0.06)
av. cognitive ability			-0.00	-0.00
			(0.00)	(0.00)
prop. parental degree			-0.06	-0.06
			(0.07)	(0.07)
friends: av. univ. aspirations * migr.b.				-0.01
				(0.05)
Num. obs.	7198	7198	7198	7198
Log Likelihood	-4099.35	-3217.30	-3191.14	-3191.13

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

All models control for survey country and sampling stratum.

Abbreviations: migr.b. = migration background.

Table 3.3: Logistic regression predicting individual university aspirations at wave 2 (reporting average marginal effects, standard errors clustered on class level). Friendship group as the social milieu.

	Model 1	Model 2	Model 3	Model 4
fr. group: av. univ. aspirations	0.39***	0.17***	0.12***	0.12***
	(0.02)	(0.02)	(0.02)	(0.02)
individual univ. aspirations, w1		0.39***	0.38***	0.38***
		(0.01)	(0.01)	(0.01)
gender (female)		0.06***	0.06***	0.06***
		(0.01)	(0.01)	(0.01)
age (in months)		-0.01***	-0.00***	-0.00***
		(0.00)	(0.00)	(0.00)
migration background (vs. native)		0.10***	0.09***	0.09***
		(0.02)	(0.02)	(0.02)
cognitive ability index		0.01***	0.01***	0.01***
		(0.00)	(0.00)	(0.00)
at least one parent has a univ. degree		0.07***	0.06***	0.06***
		(0.02)	(0.02)	(0.02)
parental univ. aspirations		0.17***	0.16***	0.16***
		(0.02)	(0.02)	(0.02)
social milieu: friendship group				
av. cognitive ability		0.01***	0.01***	0.01***
		(0.00)	(0.00)	(0.00)
prop. parental degree		0.08**	0.06*	0.06*
		(0.02)	(0.02)	(0.02)
prop. natives			-0.03	-0.03
			(0.02)	(0.02)
school class				
av. univ. aspirations			0.26***	0.26***
			(0.05)	(0.05)
prop. natives			0.05	0.05
			(0.04)	(0.04)
av. cognitive ability			-0.00	-0.00
			(0.00)	(0.00)
prop. parental degree			0.01	0.01
			(0.06)	(0.06)
fr.group: av. univ. aspirations * migr.b.				0.01
				(0.04)
Num. obs.	7198	7198	7198	7198
Log Likelihood	-4156.72	-3222.64	-3195.35	-3195.34

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

All models control for survey country and sampling stratum.

Abbreviations: fr.group = friendship group, migr.b. = migration background

Table 3.4: First-difference models with individual-level fixed effects, predicting change in individual university aspirations between wave 1 and 2

	Model 1	Model 2
Change of aspirations of friends	0.07***	0.07***
	(0.02)	(0.02)
Change of aspirations of class		0.02
		(0.04)
Intercept	0.07***	0.06***
	(0.01)	(0.01)
Adj. R^2	0.00	0.00
N	7147	7147

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Class-room level average aspirations were computed leaving out individual aspirations and friends' aspirations.

group, respectively[1].

Table 3.4 shows the results from the first-difference models with individual-level fixed effects, predicting change in individual academic aspirations by change in the predictor variables[2] The same setup is applied separately for students with a migration background and for natives, see Table 3.5.

Hypothesis 1, stating that educational aspirations are clustered in social milieus, is supported. The logistic regression models show clustering in the social milieu that goes beyond individual characteristics (see Tables 3.2 and 3.3). This applies equally to both dimensions of social milieu: Both the models focusing on friends and the models focusing on friendship groups show the same substantive results, while even the coefficients are similar. In addition, the first difference models show that change in the aspirations of the social milieu has an effect, whereas change in the aspirations on the classroom level does not (see Table 3.4). It is important to note that we find these effects while controlling for a rich amount of relevant variables.

Our second hypothesis, stating that there is no substantial difference in the clustering of educational aspirations on the level of friends and on the level of the friendship

[1]Results based on non-imputed data can be found in Tables B.2 and B.3 in the appendix.

[2]Since all independent variables included in these models are based on network items and hence could not be imputed, these results are based on non-imputed data.

group, is supported. The differences in the coefficients are not very high (see Tables 3.2 and 3.3); while in the simple model friends' aspirations have an average marginal effect of 0.49 on individual aspirations, and average aspirations in the friendship group has one of 0.39, the effect sizes converge to virtually the same coefficient in Models 3 and 4.

Hypothesis 3, stating that the social milieu remains important even when we consider contextual effects, is supported. Contextual effects on the class-level are important dimensions in which academic aspirations are clustered. This is visible in the effect sizes and significance levels of the predictors in the regression models, as well as in the decreasing effect sizes of the aspirations in the social milieu (see Tables 3.2 and 3.3). The fact that the effect size of academic aspirations on the class-level is twice as high indicates that baseline clustering on the class-level is substantial. At the same time, other contextual characteristics, such as average cognitive ability and parental degree on the class-level are not significant, while they are on the level of the social milieu. Nevertheless, the social milieu, both as in friends and as in friendship groups, remains significant with a relatively strong effect size upon introduction of class-level characteristics. An even stronger tendency can be seen in the first-difference model, which indicates that individuals change their aspirations more in line with their social milieu than with their classrooms.

We find some evidence that supports hypothesis 4, which states that clustering is less pronounced for adolescents with a migration background. In the logistic regressions, we first see the expected positive effect on university aspirations of those with a migration background. They are more likely to aspire to a university degree in comparison to natives and the size of the effect does not change throughout the model build-up. This is the case in the set of models using friends as the social milieu (Table 3.2), as well as in the set of models using friendship groups (Table 3.3); furthermore, the coefficients are the same. The interaction effect of average university aspirations of others in the social milieu and migration background, however, is non-significant with a negligible

Table 3.5: First-difference models with individual-level fixed effects, predicting change in individual university aspirations between wave 1 and 2, separately for natives and for migrants

	Migrants		*Natives*	
	Model 1	Model 2	Model 3	Model 4
Change of aspirations of friends	0.03	0.03	0.10***	0.10***
	(0.03)	(0.03)	(0.02)	(0.02)
Change of aspirations of class		-0.02		0.03
		(0.08)		(0.06)
Intercept	0.06***	0.06***	0.07***	0.07***
	(0.01)	(0.01)	(0.01)	(0.01)
Adj. R^2	0.00	0.00	0.00	0.00
N	2418	2418	4729	4729

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Class-room level average aspirations were computed leaving out individual aspirations and friends' aspirations.

effect size. Thus, we cannot conclude that clustering based on average university aspirations of the social milieu varies by migration background. Nevertheless, we do find some evidence that clustering is indeed less pronounced for youths with a migration background in the fixed-effect models. Table 3.5 show the first-difference analysis on two sub-samples of the data, split up by migration background. As mentioned in the section on descriptive results, proportions and directions of change of aspirations are virtually the same between natives and ethnics. Yet, we see that the significant effect discussed above, signifying that individuals change their aspirations with their friends, is only present in the sub-sample of natives. This points to the tendency that for those with a migration background, friends' aspirations matter less.

3.4.3 Robustness checks

A number of robustness checks were carried out in the course of this analysis. We ran the analyses on different sub-samples of the full data: we tested our hypotheses separately for each survey country, and ran the analyses by school track in Germany and the Netherlands. The country-level analyses showed differences (see Tables B.4 to B.9 in the Appendix): The results are clearest for the Swedish sub-sample. For

the German and Dutch sub-sample, in neither version of the social milieu, does the classroom context have a significant effect on individual university aspirations. There are two reasons why this could be the case. First, the differences could be due to lower statistical power in the German and in the Dutch data, since more cases are lost in comparison to the Swedish data. Second, it could be due to the fact, that in Germany and in the Netherlands, the educational system is tracked, and that there are fewer differences in aspirations on the class-level. Nevertheless, the pooled data (weighted) shows overall evidence for our hypotheses, as discussed, and we are confident that we find overall support for our assumed mechanisms. The analysis by school track did not show any additional differences.

As mentioned in the data section, we furthermore carried out all analyses with multinomial logistic regressions and an ordered outcome variable for the German part of the data. Even though the distribution of the original variable was skewed a lot more in the other two countries, we did the same for the Dutch and the Swedish part of the data. We furthermore checked robustness through different variable operationalisations. We used language ability instead of cognitive ability, and used the mode instead of the mean university aspirations in the social milieu and the classroom. All of these robustness checks resulted in the same substantive conclusions.

One plausible critique to the analytical strategy is that hypothesised mechanisms regarding migration background vary by migrant generation. It was attempted to split up the migration background variable by migrant generation, however, this resulted in a low n in some cases, also due to the many cases that are lost due to item non-response on the network variables, yielding non-significant results. It is however likely that the differences by migration background that are assumed here are in fact more complicated than the analysis can honour.

3.5 Discussion and conclusion

In this paper, we tested whether we can detect clustering of educational aspirations within school classes, i.e. on the peer level, and whether this differs between native and ethnic students. In our analyses, we find evidence that clustering and homophily

regarding educational aspirations is present on a level between the individual and the school class, i.e. on the peer level, or on the level of the social milieu, as we term in in this study. While initial sorting into classrooms does happen based on socio-economic status, cognitive ability etc., students in classrooms cluster further based on other factors, such as educational aspirations. This can be due to befriending those that are similar, or an increase of similarity of friends through socialisation processes. When it comes to ethnic students, our results are more mixed. Based on our analyses, we only found some evidence that clustering for migrants follows a different dynamic than for natives. This is likely due to the fact that we did not include various aspects in our analyses which are likely to play a role, for example differentiating between inter- and intra-ethnic connections, a finer definition of the ethnic category into countries of origin, migration generation, or measures of segregation and ethnic composition in the classroom. Our aim was to provide an initial exploration of this dynamic, and our results indicate a promising alley for future research that takes this issues into account more carefully.

In our analysis, we show the potential of social network data by conceptualising the social milieu in two ways: as friends and as friendship groups. We did not expect differences in these dimensions of clustering, and indeed, many of the coefficients are the same, and both analyses yield the same conclusions. Considering that the construction of friendship groups using the HCCA is time-consuming and computationally elaborate, this means that a friendship group construction like this is not always necessary. Further tests should be carried out to support this, but the strong similarity of the results underlines the impression that the utilisation of network data in the simple way that we are following in our analysis is likely to suffice. Thus, including friends and their characteristics in peer research is relatively straightforward, and it is a more accurate measure of the peer dimension than classroom-level aggregates. Lastly, this also points to the need for the collection of more large-scale social network data, since availability is still scarce.

We find evidence for clustering of educational aspirations in this paper, however, research on peer effects is usually interested in an *effect*, for example the influence of

a peer's attitude to the individual's attitude. Our finding of homophily in educational aspirations among friends and friendship groups can, however, come about through two mechanisms: the educational aspirations of one influencing the aspirations of another (peer influence), and students befriending those with the same aspiration (peer selection). While longitudinal social network models are able to disentangle those effects (Steglich et al., 2010), these analyses are comparatively less available and accessible, which can be explained by computationally demanding analyses, strict data requirements, and interpretations that are different from those of usual regression analyses. While there are more studies in social psychology that use this framework (e.g. Wölfer et al., 2015), this is less so the case in sociology. In the area of educational research, there are only a few studies that use social network data (e.g. Babcock, 2008) and even fewer that use longitudinal social network analysis (e.g. Lomi et al., 2011). To our knowledge, none of them are explicitly considering educational aspirations.

Since we cannot disentangle selection from influence effects in our analysis, we focus on the extent of clustering on the levels investigated, and come to two important conclusions. First, peers matter in the building and reshaping of educational aspirations of adolescents. When considering this finding in the frame of social stratification, we can conclude that peer effects and educational aspirations are a dimension which has received relatively little attention in the past, while it is likely to have the potential of affecting educational success beyond the social origin, which is so often assumed to be vital in this regard. Second, studies in this area should not be constrained to use classroom-based aggregates, because they mask a finer clustering on the classroom-level. Since those clusters, the social milieus, are the dimensions in which we can assume social influence to be strongest, using social network data is theoretically accurate as well as methodologically promising. Social network data are a well-suited source to measure this dimension, even when not applied in the frame of social network analysis. Our paper demonstrates the validity of two different approaches for measuring these clusters, either as direct friendship nominations or as friendship groups.

In conclusion, our study finds clustering of educational aspirations on the peer level that goes beyond clustering on the classroom-level. We furthermore find initial evidence

of different dynamics for ethnic minorities – this, however, has to be further explored in future research. By accounting for the immediate social surrounding in our analysis through the utilisation of social network data, we show the promise of identifying a social mileu in which peer effects are more theoretically convincing than they are at class-room level, as well as the potential to follow up on our results with more elaborate techniques of social network analysis.

Chapter 4

The social STEM pipeline

Boys and girls tend to prefer different subjects in school; this has implications for differential educational and occupational careers. Especially the tendency for boys to develop stronger preferences for Science, Technology, Engineering, and Maths (STEM) subjects than girls do over their school years is assumed to contribute to macro-level gender differences in income and social status. Relying on large-scale panel data on adolescents (the Swedish subsample of the *CILS4EU* data set, age 14/15), we find a 21% drop over a year in the number of girls who choose a STEM subject as their favourite subject (from 19% to 15%), whereas this is only 5% for boys (from 21% to 20%). Through the utilisation of social network data (218 classrooms, 4,998 students, and a total of 27,428 friendship relations at two time points) and the application of the newly developed random-coefficient multilevel stochastic actor-oriented models (multilevel SAOMs), we investigate how peer effects shape these changes. We simultaneously test two substantially different mechanisms: adjusting subject preferences to those of one's friends (friend influence) and following gendered behaviour specific to the classroom (peer exposure). We find strong evidence for both of these tendencies. Amplifying pre-existing gender preferences, these mechanisms contribute to the dramatic widening of the gender gap in STEM preferences in our sample. We conclude that explanations

This chapter is based on the paper "The social pipeline: How friend influence and peer exposure widen the STEM gender gap", that is currently being revised for publication at *Sociology of Education*. It is co-authored with Christoph Stadtfeld and Zsófia Boda, Chair of Social Networks, ETH Zürich.

of why more girls leave STEM-career-paths benefit from considering detailed social processes in the classroom.

4.1 Introduction

Research on horizontal sex segregation in the labour market, i.e. the tendency for men and women to work in different fields of occupation, has documented a process of girls and women to drop out of a Science, Technology, Engineering, and Maths (STEM) career path at a higher rate than men throughout the life course. That girls tend to increasingly like subjects other than STEM over their school career is one example of this phenomenon, and is a puzzle that has been investigated by past research through a variety of factors, such as socialisation. While early life socialisation happens mostly within the family, the role of peers becomes increasingly important when growing up: during adolescence, interaction and exchange with friends, especially in school, is considered to be a crucial setting in which socialisation happens (Osterman, 2000). Adolescence is an important period of life, not only because of peer socialisation, which has implications for life-long attitudes, norms, and values, but also because it is the time in which life choices regarding one's career are made. Those can have long-lasting effects on life outcomes, such as income and social status. Therefore, many studies have attempted to investigate peer effects and social influence among adolescents in the school context.

In this paper, we consider two main ways in which peers matter for subject preferences in high school: friend influence and peer exposure. Friend influence refers to the tendency that individuals develop preferences for subjects that their friends like. This type of peer effect, thus, focuses on direct influence from friends, i.e. those in the classroom community that are closest to the individual, and who therefore will have a higher impact relatively to others. This is in line with previous research on friend selection and influence on a variety of behavioural outcomes (Veenstra et al., 2013). Peer exposure expresses a tendency that in classrooms in which girls generally like STEM subjects, more female students develop similar preferences. This type of peer effect captures the classroom context regarding the preferences of other girls, and is advancing previous

research on classroom context, which made use of classroom-level aggregates and their association with individual outcomes. Since differential STEM preferences are assumed to be at least partly affected by the extent to which individuals adhere to gender roles in the group context, the extent to which girls prefer STEM subjects in a class is likely affecting those preferences as well.

We contribute to existing research in several substantial ways. By using multilevel network models on a large sample, we are able to distinguish between influence and selection effects (e.g. whether friends influence one's subject preferences, or one chooses friends with similar preferences). Thus, we can explicitly account for dynamics in the immediate social surroundings of the individual, rather than solely considering an aggregate measure on the class room or school level. At the same time, we are able to include the peer context within classrooms in our analysis as well. Lastly, to our knowledge, no researchers have attempted so far to apply social network models to investigate social influence effects on gender differences in STEM preferences.

4.2 Theory, previous research, and hypotheses

This section gives an overview of relevant theory and previous research, and develops the hypotheses guiding the research presented in this chapter.

4.2.1 Occupational gender segregation and the leaky pipeline

More boys than girls, and more men than women are interested in STEM; the differences between the two genders continuously grow during childhood and adolescence. Differences in the extent to which students prefer STEM subjects in high school develop into occupational sex segregation in adult life, i.e. the fact that men and women are often not distributed evenly across occupational hierarchies and sectors (Browne, 2006; Charles and Grusky, 2004; Inglehart and Norris, 2003).

Different forms of segregation have been documented repeatedly in the last decades, and are still persistent (Levanon and Grusky, 2016). While vertical gender segregation (i.e. men occupying higher paying jobs) has started to equalise somewhat in the last decades, horizontal gender segregation (i.e. men and women working in different sec-

tors) has changed little (England et al., 2007). One sector in which vertical segregation is especially apparent is the field of STEM: a recent report published by the European Union states that in almost all European countries the proportion of female scientists and engineers in the total labour force is lower than the male proportion (European-Commission, 2016).

The relatively higher rates with which females drop out from STEM educational tracks and occupations, thus creating and widening the gender gap, is often metaphorically described as a leaky pipeline. STEM career paths are often conceptualised as a pipeline, since they require a relatively strict pathway from subject electives and advanced courses in high school, to an undergraduate degree with a STEM focus, to a postgraduate degree, to a first job: Joining this trajectory at a later stage is difficult or even structurally impossible, since often specific prerequisites of previous STEM involvement have to be fulfilled. When it comes to leaving this trajectory, it has been found that women drop out ("leak out") of the pipeline at a higher rate than men, although there are no significant gender differences in performance in relevant subjects at any stages of the education system (Alper, 1993). Favourite subjects in school are part of this trajectory. Students will pick advanced courses based on this, and eventually this will inform choices in the tertiary education, which then determines trajectories in the labour market.

While the leaky pipeline model provides a life-course perspective on segregation, it only shows that women leave the STEM-pipeline at a higher rate than men, and it does not offer an explanation. There are several approaches aiming at filling this gap, covering a variety of different angles, including both micro- and macro-level approaches (e.g. England et al., 2007; Reskin et al., 1990). In this paper, we are focusing specifically on individual-level factors that are particularly relevant to the educational context.

As mentioned above, there are hardly any gender differences when it comes to maths and analytical skills in school. However, girls are more likely to have both high maths and verbal skills, which can be interpreted as them having more choices when it comes to careers (Wang et al., 2013). Following this argument, more women are expected to choose a path related to their non-STEM skills, even if just by chance, but potentially

also due to avoidance of competition. The fact that girls outperform boys so much in verbal skills, and therefore have their comparative advantages in subjects such as languages, also means that they may prefer these subjects even if they do quite well in STEM subjects because their relative position is worse there (Jonsson, 1999).

Another important approach focuses on the extent to which adolescents incorporate gender roles that relate to skills and aptitude in society into their identity (Akerlof and Kranton, 2000; Sinclair and Carlsson, 2013), such as the view that women should be more caring and nurturing and men should be providing for their family (Kroska, 2007). While this does not include a hierarchical notion per se, it leads to a narrative that women are particularly suited for subject areas that lead to jobs that tend to be in low-pay and social status, and that men are particularly suited for highly prestigious and highly paid jobs. Cross-national variations in STEM gaps, and the fact that we find gender gaps in educational and occupational choices in the most egalitarian countries (McDaniel, 2016) as well as individual level differences suggest that these gender norms are not universally determinative, but depend on the context.

4.2.2 Peer effects and gender roles

To what extent these norms on perceived gender differences are salient in individuals' choices depends on a variety of factors. Presumptions about male and female characteristics are disseminated and reinforced through popular culture, the media, and through individual experiences and events. In these situations, evidence that is consistent with pre-existing stereotypes tends to be remembered, and counter-examples tend to be ignored or forgotten (Correll, 2001). These dynamics are also shaped through the relationship to significant others, i.e. peers or family members: social interactions can implicitly or explicitly support or reject gender stereotypes (Levanon and Grusky, 2016). While at an early age the family is the most important frame of reference, peers become more relevant during adolescence (Osterman, 2000; Youniss and Smollar, 1985), which is a time in which already existing preferences are reframed and when young people make decisions about their future. This phase of the life is thus of particular interest when it comes to STEM preferences, and the way peers adhere to gender

roles or sanction atypical behaviour, is likely to have an impact on individual behaviour.

Measuring peer effects, however, comes with conceptual challenges. Most importantly, deciding on who the relevant peers are is not trivial. It can be argued that the most important peers are probably to be found in their school class, since youths spend the majority of their time in this context and interact with others of the same age and in the same life situation. Thus, a lot of studies conceptualise the peer effect as an aggregate of the characteristics of the classroom context (Legewie and DiPrete, 2014; Rjosk et al., 2014; van Ewijk and Sleegers, 2010). Although these studies theoretically consider that individuals are influenced by significant others on the micro level, the methodological approach often only considers the classroom context through the utilisation of aggregates. This is often done due to the lack of data on within-classroom relations. Friendship network data allow researchers to explicitly consider particular peers. Indeed, there has been some convincing evidence showing that "some are more peer than others", suggesting to focus on peers "subjectively meaningful to individual" when researching peer effects (Lomi et al., 2011). In our study, we focus on two kinds of peers who are subjectively relevant to the individual: friends and others in the classroom.

4.2.3 Friend influence effects

Friends have been long considered to be of relevance in theories of educational success. The influential Wisconsin model of status attainment assumes influences from expectations of "significant" peers on the individual, thereby implicitly differentiating close friends from classmates (Sewell et al., 1969). Homophily, the tendency for friends to be similar in multiple regards, e.g. in their gender, age, attitudes, and cultural taste, is widely documented (McPherson et al., 2001). This can be the result of friend selection, i.e. the tendency of befriending others who are similar, but, regarding changeable characteristics, can be evidence for social influence. A lot of research on friendship networks has documented the existence of this tendency regarding a variety of outcomes (Cheadle et al., 2015; DeLay et al., 2016). When it comes to favourite subjects in school, it is likely that similar dynamics apply. Youths are likely to value their friends'

opinion, and are likely to interact with them on an almost daily basis. They likely select their friends based on a variety of important and often stable characteristics, like their gender, their interests and their cognitive skills. Friends are likely to select or influence each other in changeable characteristics too, so that they are likely to share general attitudes and tastes, are thus used to supporting the same ideas, and be attracted to the same challenges and contents that a school subject offers. We thus hypothesise:

Hypothesis 1: Students adopt their friends' subject preferences over time. (Friend influence effect)

Given that boys and girls show different subject preferences in school, and their friendship patterns tend to be somewhat different (Rose and Rudolph, 2006), it is possible that they are also subject to different peer effects. Therefore, we will test whether there are gender differences to the extent to which boys and girls are susceptible to friend influence.

4.2.4 Peer exposure effects

While there is evidence for friends to be the more determinative frame of reference in socialisation in comparison to all others in the same context (such as in the classroom), effects of other relevant peers are also important to consider. When it comes to the politics of gender norms, behaviour of others from the same gender can be particularly crucial. The reasoning for this is rooted in the nature of norms themselves, as well as in previous research on exposure to gender-typed behaviour.

First, research in developmental psychology has found that adolescents are sensitive to behaviour of others in the same social category, i.e. to others of the same gender and the same age. The extent of exposure to peers of the same gender is positively correlated with the extent of gender-typed behaviour individuals exhibit (Martin and Fabes, 2001). Second, social norms are highly dependent on the context in which they are applied (Bicchieri, 2006). For individual behaviour they can be guiding or even restrictive, due to sanctions that can be applied. Importantly, a norm is only salient

when the vast majority of others adhere to it. There are several examples for norms guiding human behaviour, such as shaking hands after a tennis match: A player not conforming to this would experience social sanctions. This is only the case, though, since everyone else who plays tennis adheres to this norm, unlike in other kinds of sport. Indeed, results from behavioural experiments suggest a dynamic interplay of the social surrounding and the existence of norms in the particular social context (Efferson et al., 2016; Block et al., 2018).

Thus, it is likely that the extent to which gender norms are upheld by others in the classroom matters for the individual, even if these others are not his or her friends, since they are in a position to exert sanctions. Adolescence is a vulnerable phase, which leads to the particular need to belong to the school community (Osterman, 2000). Since adolescents tend to adhere to gender roles, it is likely that non-conforming to gender-typed behaviour and preferences is sanctioned by classmates. The costs of non-conforming, however, will become lower (or even disappear) the more individuals exhibit non-gender-typed preferences: either since this social norm becomes more and more deconstructed, or because sanctions, such as excluding individuals from the community, are not effective as otherwise isolated individuals will have each other.

We thus propose that gender roles are more salient the more girls adhere and support them; the more girls behave in a non-gender-conform way, the less salient this norm will become, and hence, the less influence it will have on others. Thus, we expect STEM preference of girls to be more stable and more prevalent in classrooms in which there is an overall tendency for girls to like STEM.

Hypothesis 2: Exposure to gender atypical subject preferences of girls in a classroom influences other girls to adapt these preferences. (Peer exposure effect)

It is furthermore plausible, that peer exposure effects on STEM preferences work differently for same or opposite sex friends. To address this, the effect of male STEM preferences in a classroom on girls is tested as well.

We assume that the two types of peer effects exist at the same time. Given that

during adolescence pupils mostly have same-gender friends, the two mechanisms often overlap. Modelling them together, we are able to disentangle them from one another.

4.2.5 The challenge of studying peer effects

In this study, we propose to consider both friend influence and peer exposure effects. Both conceptualisations have been researched in the past, and both have methodological or conceptual shortcomings. We aim at rectifying this, by including both types of peer effects in our analysis at the same time, and by utilising the novel approach of random-coefficient multilevel SAOMs.

First, as argued above, it is more exact to use friendship network data than classroom level aggregates when researching peer effects. When considering friends to understand their effect, however, it is important to consider how friendships change over time. This is necessary to distinguish a peer influence effect, i.e. changing behaviour because of an existing friendship, from a peer selection effect, i.e. befriending someone who is similar. Regression-based analyses considering the effect of friends (Carbonaro and Workman, 2016) can associate friends' characteristics with those of the individual over time, but they cannot account for these different friendship processes. SAOMs that apply longitudinal social network analysis, however, are able to do so (Steglich et al., 2010).

Second, until recently it has been challenging to study multiple groups (for example school classes, as in our study) in the SAOM framework, which is necessary to draw generalisable conclusions. While a variety of methods to analyse multiple groups (e.g. meta-analysis, multi-group analysis; see Ripley et al. 2017) have already been available for a while, they often proved suboptimal, due to the lack of statistical power (Stadtfeld et al., 2018), or strong statistical assumptions. The recently developed random coefficient multilevel network models are able to rectify those shortcomings. They simultaneously estimate the entire group of networks in a Bayesian framework, while being able to take between-group differences into account. Thus, they do not struggle with model convergence and statistical power, and at the same time are able to account for multilevel dynamics (Snijders and Koskinen, 2017).

Our study contributes to previous research in four ways. First, we are able to capture the theoretically relevant peer dimension, i.e. friends, through the utilisation of social network data. Second, we are simultaneously considering exposure to classmates as well, in line with previous research using classroom-level aggregates. Thus, we are able to show friend influence net of peer exposure, and vice versa. Third, we are applying SAOMs to disentangle selection and influence effects. Fourth, we are using random coefficient multilevel network models to analyse a large group of networks from a large and generalisable data set, being able to account for multilevel dynamics.

4.3 Data and Methods

This section introduces the data and methods used in the analysis.

4.3.1 Data

We use the Swedish part of the first two waves of the Children of Immigrants Longitudinal Survey in Four European Countries (*CILS4EU*) dataset (Kalter et al., 2014, 2015). Within the frame of this project, which is funded by several European research councils (NORFACE), adolescents in England, Germany, the Netherlands and Sweden were followed over three waves in three years, from age 14/15 on. Parental and teacher surveys were carried out, too. The sampling design took a two-step cluster approach, first, selecting schools (oversampling schools with a high proportion of immigrant students), and then randomly drawing two classes within this school. In Sweden, 5,025 students from 251 classes in 129 schools participated in the survey (Kalter et al., 2014, 2015). Social network data were collected as part of the youth survey, and contain information on a variety of social networks, including friendship networks.

We focus on the Swedish part of the data for conceptual as well as technical reasons. Conceptually, the non-tracked education system of Sweden implies that the students in our sample structurally have the same choices and prospects, while in the Netherlands and Germany, where the education system is tracked, this is not the case. Additionally, in the Netherlands, longitudinal networks analysis is almost impossible, due to a sorting into field of study that takes place after wave 1, which involves a restructuring

of classes. England has a non-tracked education system as well, but there are some inconsistencies in the collection of the network data (Kruse and Jacob, 2014). With Sweden, we are focusing on a country that is considered to be one of the more gender-equal European countries, due to, for example, governmental policies aiming at equal effects of child-bearing on labour market chances (Duvander and Andersson, 2006) and high post-materialist value orientation (Inglehart, 2008). Nevertheless, the gender differences in the proportion of women scientists and engineers is still substantial, though less pronounced than in other European countries (EuropeanCommission, 2016). This points to the limited unconditional influence of a gender-egalitarian ideology on educational and occupational choices, and it suggests that any evidence for gendered pathways to STEM that we find in Sweden is also likely to apply in less gender-equal countries.

4.3.2 Variables

We measure preferences for STEM subjects as favourite subjects. The survey item that this measure is built on asks "What is your favourite subject at school?"; multiple answers were possible. Most respondents answered with one subject, and nobody indicated more than two subjects. The answers were grouped into 17 categories: Art, Biology, Chemistry, English, Handicrafts, History, Home economics, Mathematics, Music, Natural Science, Other foreign languages, Physical education, Physics, Religious Studies, Social studies, Swedish, and Technology. Of these, we consider Chemistry, Mathematics, Natural Science, Physics, and Technology to be STEM-subjects. Biology is not included in this list, following a strand of literature that considers Biology to be a "soft science", thus having different implications (Gabay-Egozi et al., 2014; Smith, 2011).

Gender (female or male) is defined based on self-reports. We furthermore control for two exogenous factors that could be related to STEM preferences: socio-economic status and cognitive ability. Socio-economic status is measured the as highest form of parental educational attainment in four categories: no school leaving certificate, degree from lower secondary school, degree from upper secondary school, and university

Table 4.1: Descriptive statistics and number of cases of analytical sample; covariate variables.

	mean	SD	median	min	max	N
Cognitive ability test	17.54	4.94	18	0	27	4189
Gender: female						2518
Gender: male						2460
Students having a STEM-favourite subject, wave 1						862
Students having a STEM-favourite subject, wave 2						687
Parental education: no school leaving certificate						173
degree from lower secondary school						164
degree from upper secondary school						1480
university degree						2055

degree. This information is drawn from the parental questionnaire; where possible, missings due to parental non-participation were replaced by information from the youth questionnaire. Cognitive ability is based on a language-free test, that is part of the *CILS4EU* data. Descriptive statistics on all variables in the analytical sample can be found in Table 4.1, for the full sample in Table C.1 in the Appendix.

4.3.3 Friendship and favourite subject networks

Friendship networks are measured through the survey item "Who are your best friends in class?"; it was possible to nominate five others. Table 4.2 presents network level descriptives of friendship and subject affiliation networks, respectively. The descriptives as well as all subsequent analyses are based on a sub sample of 218 school classes. 33 of the 251 school classes were excluded as they were related to estimation issues in the *RSiena* estimation routine. Descriptive comparisons indicated that the excluded networks are not very different from the 218 classes in the sample but tend to be somewhat smaller. Details are presented in Table C.2 in the Appendix. Each table includes the range, the mean, and the standard deviation of a number of indexes. The mean number of students per class is 22.9 with the smallest and largest class having 9 and 34 students. Network densities are defined by the number of relations over the

Table 4.2: Network descriptives, groups included in the analysis.

	min	max	mean	sd
friendship networks				
size	9.00	34.00	22.86	3.67
density	0.00	0.24	0.13	0.04
degrees	0.00	5.00	2.74	1.85
avgDegree	0.00	4.70	2.71	0.73
clustering	0.25	0.88	0.57	0.11
jaccard	0.00	0.58	0.33	0.11
same gender ties, wave 1	0.54	1.00	0.90	0.08
same gender ties, wave 2	0.51	1.00	0.87	0.10
subject networks				
density	0.03	0.08	0.05	0.01
degrees	0.00	2.00	0.89	0.42
avgDegree	0.48	1.39	0.89	0.13
jaccard	0.05	0.56	0.27	0.09

maximum possible number of relations (if, for example, everybody was friends with everybody else in a particular school class). On average, the density is 13% for the friendship networks (individuals indicate to be friends with 13% of their class mates) and 5% for the subject affiliation networks. Students only indicate 0.9 favourite subjects on average, but 2.7 friends. Friendship networks have an average clustering coefficient of 0.57. This indicates that 57% of all two paths (i nominates j, j nominates h) are transitively closed (i also nominates h). Both networks are on average similarly stable which is indicated by average Jaccard coefficients of 33% (friendship) and 27% (subject affiliation). The Jaccard coefficient is the proportion of ties that are stable (thus exist in both data collection waves) over the number of ties that exist in either or both data collection waves. It ranges from zero to one, where one indicates that the network does not change at all. Some friendship networks have a Jaccard coefficient of zero, which is due the fact that in some classes no friendship information was collected in the second wave and the networks thus appear empty; since there is information in the favourite subject networks, however, these classes still carry usable information for our analysis. The classrooms also show strong gender homophily, which is typical of adolescent friendship networks: 90% of friendships are same-gender in wave 1, and 87% in wave 2.

4.3.4 Method and analytical strategy

For our analysis, we use Stochastic Actor-Oriented Models (SAOMs, Snijders, 2001, 2011; Steglich et al., 2010). These models require longitudinal (panel) data and use simulations to infer social mechanisms behind the observed tie changes in the network. Network changes are represented as a sequences of many small changes, such as in an agent-based simulation model. In each step, one randomly selected actor has a chance to change (create or terminate) an outgoing tie (or to keep the network unchanged). The first wave of data serves as a starting point for modelling the second wave. During the simulation, actors make choices about whom they want to be connected to based on theoretically assumed effects: these serve as independent variables in the model.

In the SAOM framework, it is possible to model the co-evolution of one-mode and two-mode networks (Snijders et al., 2013; Stadtfeld et al., 2016). In a one-mode network, all nodes are of the same type (e.g. students), and ties can exist between pairs of these nodes. In a two-mode network, relations connect nodes of two different node sets (here: students and school subjects). In this paper, we model the co-evolution of one-mode networks where the nodes are students and ties represent friendships between them, and two-mode networks where nodes are both students and all possible subjects, and ties represent that a given student considers the given subject as his or her favourite. Visualisations of several sample classrooms are presented and discussed in the section on descriptive results.

When analysing the co-evolution of a one-mode and a two-mode network, students are assumed to make decisions about their friendship ties and about their favourite subjects through time. Friendship evolution and favourite-subject evolution are modelled as two dependent variables, based on two separate sets of independent variables. In this way, it is possible to express interdependencies between the two levels and to include the effect of someone's favourite subjects on his or her friendships, and that of one's friendships on his or her favourite subjects. This has two major advantages. First, it allows us to distinguish between selection (friendship choice based on favourite subject) and influence (favourite subject choice based on friendship) processes. Second, it allows us to separate both selection and influence effects from endogenous effects of

friendship or favourite subject choices. Examples of endogenous effects are transitive closure in the friendship network (friends of friends being friends) and Matthew effects in the subject network: School subjects might become even more popular in the future purely due to their past popularity. We can further control for individual mechanisms, for example, the tendency that students with higher cognitive skills are more likely to choose specific subjects. While traditional regression-based methods consider observations (i.e., social ties or favourite subjects) to be independent, social network models capture and statistically model these dependencies, controlling for them when estimating the rest of the parameters.

A methodological challenge in education research is its multilevel nature. Researchers need to aggregate classroom-level results to global estimates, while taking into account effects on the classroom level, for example, the behaviour of peers. In our network approach, we additionally need to take into account the observational dependence that originates from network structures. Thereby, we can, for example, express the effect of the behaviour of someone's friends (net of classroom peer effects). To capture the complex multilevel structures and network dependencies in our sample, we apply a new methodological addition to SAOMS that allows the fitting of random-coefficient multi-level network models. These SAOMs can be imagined as being analogous to the hierarchical linear model for multilevel analysis in the non-network framework (Snijders and Koskinen, 2017; Ripley et al., 2018) but additionally, they allow to test for network dependencies. The method makes use of a Bayesian estimation technique, which has the additional advantage of providing increased statistical power in cases when model specifications are too complex for the sizes of individual classrooms (which applies to many longitudinal selection-influence models, (see Stadtfeld et al., 2018). Details on the estimation procedure are provided in section C in the Appendix.

Our analyses have been carried out in *R*, version 3.2.2, using the package *RSienaTest*, version 1.2-2 on a large-scale computer cluster.

4.3.5 Model specifications

To test how friends influence each other regarding subject choice, we include an effect that models the tendency of those who are connected in the friendship network to make the same subject choices in the favourite subject network: *Friend influence*[1]. Friend influence here is to be understood as the tendency for two friends to start liking or keep liking the same subject between wave 1 and 2. Since this is a multinomial choice, there is either agreement on a favourite subject or not. We control for subject homophily which is specified as the tendency to be friends with others who have the same favourite subject: *Shared favourite subject of ego and alter*. The effects are similar, but in case of friend influence the dependent variable is the subject network, in case of subject homophily the dependent variable it the friendship network. Since we are interested in whether these tendencies vary by gender, we are interacting the two effects with the gender covariate *female ego*.

To test the peer exposure effect, we include a number of main effects and their interactions in the favourite subject network part of the model. The effect that directly tests our hypothesis is the three-way interaction of

1. *subject popularity among females* (are subjects that are liked by girls in general more popular?)
2. the subject covariate *STEM subject* (are STEM subjects more popular in general?)
3. the actor-covariate *female ego* (are girls in general more likely to have a favourite subject?)

Thus, this interaction tests whether girls like STEM subjects more if it they are generally more popular among girls, in comparison to whether they are not generally more popular among girls. We furthermore include the respective three two-way interactions and the three main effects into the model. To test for opposite sex peer exposure, we

[1]Names in italics are used in the results tables. A full overview of all effects included in the analysis, as well as the internal *RSiena* shortnames, is given in Tables C.3 (friendship part of the model) and C.4 (favourite subject part of the model) in the Appendix.

are including a three-way interaction with subject popularity among males along with the necessary two-way interactions in an additional model.

Apart from the effects that relate to our hypotheses, we include a number of controls in our model, both for endogenous network processes and for alternative explanations that might relate to differential STEM preferences. As network controls, in the friendship network part, the following structural effects are included: reciprocity, transitive triplets, transitive reciprocated triplets, degree-related effects, as well as effects to capture gender homophily. In the favourite subject part, we include structural effects related to the degree popularity, i.e. the tendency that subjects become more popular over time since they are already popular. We furthermore control for additional cross-network effects, taking into account tendencies of friendship popularity and nomination behaviour to affect favourite subject nomination behaviour. To account for individual characteristics that might influence the extent to which individuals like STEM subjects, we include actor covariates for cognitive ability and parental socio-economic status (SES) in the favourite subject network part, and interact them with the subject covariate STEM.

Based on visual inspection, the model shows good convergence. Following Gelman et al. (2013) and Snijders and Koskinen (2017), a convergence $\hat{R}$ was also calculated, indicating good overall convergence for our model.

4.4 Results

Before we present results from the multilevel network models, some descriptive results on the networks and variables of interest are shown and discussed.

4.4.1 Descriptive results

Figure 4.1 shows six exemplary combined networks (friendship and subject affiliation networks) from wave 1 of the data. Individuals are shown in the centre of the plots (females as triangles, males as circles) and subjects as squares in the periphery. An arrow between two individuals indicates a friendship nomination, an arrow between an individual and a subject indicates a favourite subject nomination. From the visual

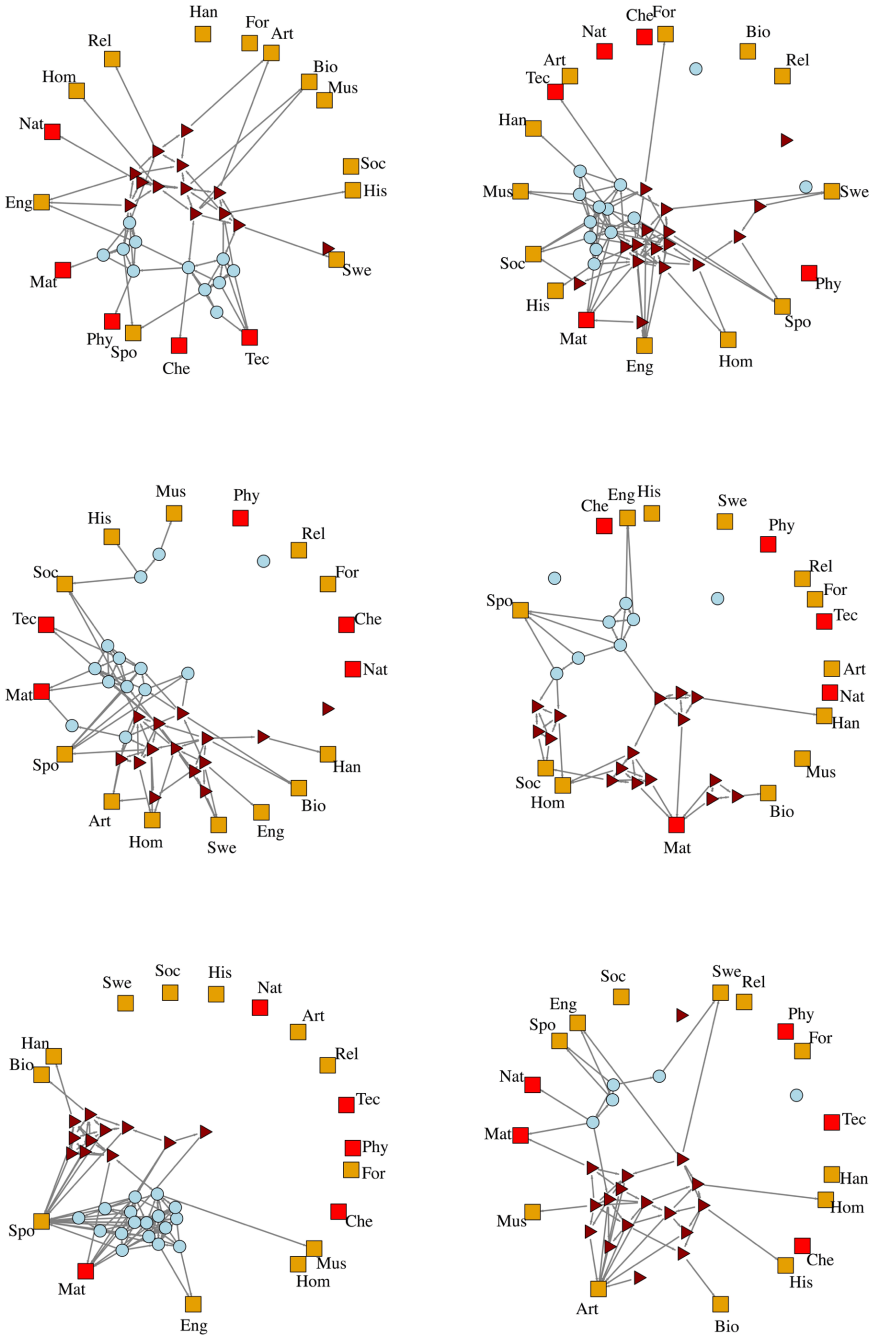

Figure 4.1: Combination of one-mode and two-mode plots, six sample classrooms. Triangles represent girls, circles boys. Subject abbreviations: Art = Art, Bio = Biology, Che = Chemistry, Eng = English, Han = Handicrafts, His = History, Hom = Home economics, Mat = Mathematics, Mus = Music, Nat = Natural Science, For = Other foreign language, Spo = Physical Education, Phy = Physics, Rel = Religious Studies, Soc = Social studies, Swe = Swedish, Tec = Technology.

representation one can expect that gender homophily (a higher tendency of friendship between same-gender individuals) and transitivity (the tendency to form small groups) are potentially important social forces that explain change in the friendship network. The subject affiliation seems to be partly shaped by a tendency towards degree centralisation, considering the large number of potential subjects in relation to number of students in the classroom.

In general, there are gender differences in favourite subjects, which are in line with our expectations, as shown by Figures C.1 and C.2 in the Appendix. Subjects like Arts, Swedish, or English are more popular among girls, while Physics or Technology are more popular among boys. Physical education is very popular in general, but is named by substantially more boys than girls. There are some variations between wave 1 and 2: Maths, for example, is equally popular among boys and girls in wave 1, but is named by more boys in wave 2.

When considering the dynamics of STEM subjects in particular, we see that STEM subjects become less popular between wave 1 and 2; this decrease in popularity is stronger for girls. In wave 1, 21% of all boys and 19% of all girls have a favourite subject that is a STEM subject. In wave 2, this applies to 20% of all boys, and 15% of all girls. In other words, between wave 1 and 2, approximately 21% of girls "leak" from the STEM pipeline, while in the same amount of time, only 5% of boys "leak". More detailed information can be found in Table C.5 in the Appendix. This descriptive finding is perfectly in line with the leaky pipeline model.

Figures 4.2 and 4.3 show how STEM preferences differ between the school classes in the data, between wave 1 and 2, respectively. The same overall picture applies: STEM preferences decrease in general, and more so for girls. Additionally, the figures show that STEM subjects are differentially popular in different classrooms, and that the extent to which this varies by gender is not uniform, either. This warrants the application of a multilevel framework.

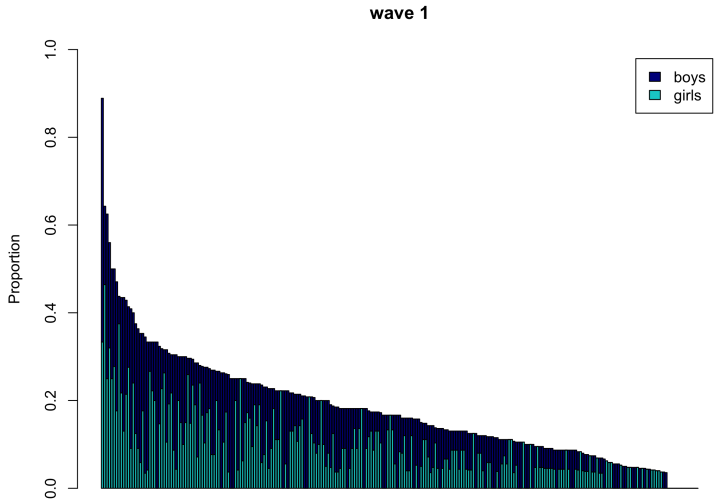

Figure 4.2: Proportion of students in each classroom that have STEM-favourite subject wave 1; proportion split up by gender.

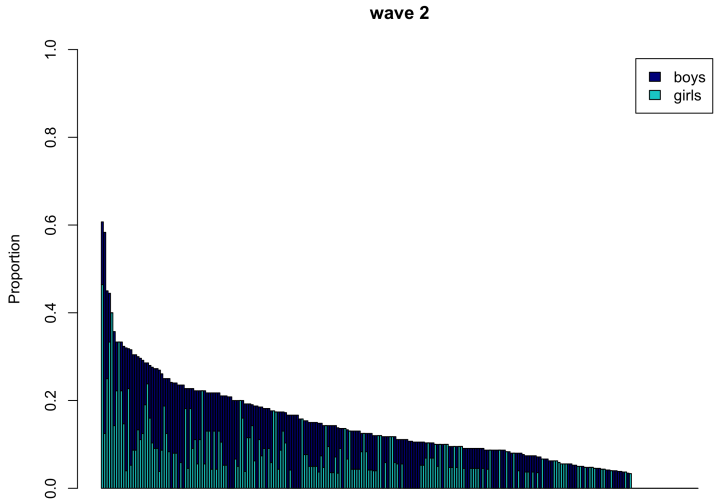

Figure 4.3: Proportion of students in each classroom that have STEM-favourite subject wave 2; proportion split up by gender.

4.4.2 Results from multilevel network models

As discussed above, we analyse peer effect on favourite subjects by modelling the co-evolution of friendships and subject choices of students. Therefore, our model consists of two parts, one that relates to friendship dynamics, and one to the dynamics in the favourite subject network. In this section, we only report the variables relevant for our hypotheses, see Table 4.3; the whole model, including the detailed model of friendship dynamics, can be found in Table C.6 in the Appendix. However, for the interpretation of each parameter in the model, it is important to consider that friendship selection processes have also been taken into account.

As stated in Hypothesis 1, we expected friends to influence each other in their favourite subjects (friend influence effect). This hypothesis is confirmed by our analysis: the friend influence variable shows that the more friends someone has who already prefer a certain subject, the more likely it is that this student also chooses that given subject as his or her favourite one ($\beta = 1.089, p < 0.001$ [2]). Since boys and girls might be susceptible to friend influence to a different extent, we also tested whether this effect is the same for both genders. We find a negative interaction effect between friend influence and gender ($\beta = -0.477, p < 0.001$), showing that girls are influenced by their peers, but significantly less so than boys.

According to Hypothesis 2, we expected that the more female students already prefer STEM subjects in a class, the more likely it will be for a girl to choose a STEM subject as her favourite one (exposure peer effect). In contrast to Hypothesis 1, we here consider all other girls in the classroom, and not just friends. While the effect of other girls could exist for all subjects to some extent, we specifically expected this for STEM, because liking STEM subjects is an expression of preferences not necessarily supported by gender norms in the outside society. This is indeed what we find: a general tendency of subject popularity among girls in the classroom affecting girls' individual preferences, but a stronger effect if the particular subject is a STEM subject. For

[2] Given that an interaction with gender is part of this model, the parameter actually refers to the differential tendency of boys to be influenced. A model without a gender interaction effect (not reported here, but to be found in Tables C.8 and C.9 (with subject dummies) in the Appendix) confirms that we find evidence for friend influence irrespective of gender.

Table 4.3: Results, random-coefficient multilevel SAOMs. Main model.

effect	β	SD	varying
Favourite subject part			
Female ego	0.643^{***}	(0.147)	-
Subject popularity among females	0.331^{*}	(0.141)	-
STEM subject	0.324	(0.272)	-
Female ego * subject popularity among females	-0.161	(0.100)	-
Female ego * STEM subject	-0.946^{***}	(0.190)	-
subject popularity among females * STEM subject	-0.018	(0.163)	-
Female ego * subject popularity among females * STEM subject	0.353^{***}	(0.139)	-
Friend influence	1.089^{***}	(0.091)	-
Female ego * friend influence	-0.477^{***}	(0.092)	-

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

This model controls for friendship selection, endogenous network effects in the friendship and the favourite subject part, as well as effects from cognitive ability and parental education on favourite subject. For full results see Table C.6 in the Appendix.

non-STEM subjects, the effect of a female student's preference on that of another can be expressed by the sum of the variables *subject popularity among females*, and its interaction with *female ego*. This sum is positive and significant ($\beta = 0.170, p < 0.001$). Therefore, there is already a small effect of preferences of other female students in the classroom on girls' choices, even without considering STEM subjects in particular. The three-way interaction variable *female ego * subject popular among females * STEM subjects* models the additional effect of the subject in question being a STEM subject; thus, expresses whether the effect of a girl on other girls is larger in the case of STEM subjects. The variable is positive and highly significant ($\beta = 0.353, p < 0.001$). This means that other girls' preferences in the classroom matter much more if these are expressed against mainstream gender norms in society. This provides strong evidence for our second hypothesis.

Beyond the effects described above, we find that STEM subjects are neither more nor less popular than other subjects in general ($\beta = 0.324, p = 0.089$). However, the negative interaction between the *STEM subject* and the *female ego* variables ($\beta = -0.946, p < 0.001$) suggests that girls, as opposed to boys, are significantly less likely to

Table 4.4: Results, random-coefficient multilevel SAOMs: Same and opposite sex peer exposure. Linear combination of relevant main and interaction effects.

effect	β	SD	varying
Favourite subject part			
Influence of boys' STEM preferences			
on boys	−0.247	(0.211)	-
on girls	−0.121	(0.312)	-
Influence of girls' STEM preferences			
on boys	0.303	(0.272)	-
on girls	0.457***	(0.100)	-

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

This model controls for friendship selection, endogenous network effects in the friendship and the favourite subject part, as well as effects from cognitive ability and parental education on favourite subject. For full results see Table C.7 in the Appendix.

select a STEM subject as their favourite subject compared to other subjects. This negative tendency for developing a new preference for a STEM-subject, and for maintaining existing STEM-preferences, is true even beyond our hypotheses on the friend influence effect, given that both hypothesised processes were analysed in the same model.

Testing for influence from opposite sex peer exposure on STEM subject preferences shows that the above-identified influence from girls' STEM preferences on other girls' STEM preferences is a unique case in all possible combinations (see Table 4.4). Boys' preferences do not have an effect on boys' or girls' preferences that is specific to STEM preferences; this does also not apply to girls' preferences on boys' preferences. Only girls' preferences for STEM have an effect on other girls' STEM preferences, with this effect being significantly different from non-STEM subject preferences.

4.4.3 Robustness checks

We carried out several robustness checks to validate our results. First, to capture differential popularity of specific subjects (for example physical education), we repeated our main analysis with subject dummies in the two-mode network part. Results can be found in Table C.10 (friendship part) and C.11 (favourite subject part) in the Appendix; there are no substantive differences to the results we are discussing.

We furthermore tested both peer effects separately, with and without subject dummies, see Tables C.12 to C.15 in the Appendix. In all models, the substantive results are the same: We find clear evidence for peer exposure and peer influence effects on subject preferences.

One plausible critique is that it matters for friend influence and peer exposure whether the individual in question has a romantic partner. While this is a conceptually interesting, it is empirically impossible to address in this study: It would require the addition of two further three-way interactions for the friend influence effect, and two four-way interactions for the peer exposure effect. Adding further three-way interactions destabilises the model, and four-way interactions are not implemented in *RSiena*. Since there is data on romantic partners available, it would be possible to follow up on this point with another research design. However, the number of individuals that have a boyfriend or girlfriend in the classroom, is very low (74 in wave 1, and 113 in wave 2, of 5,448 individuals; see Table C.16 for further details). Yet, this is likely to be most relevant to the proposed mechanisms. In total, i.e. inside or outside the classroom, 769 adolescents have a romantic partner in wave 1, and 874 in wave 2 (see Table C.17 in the Appendix for further details).

4.5 Discussion

In this paper, we aimed at understanding the often observed growing gender gap in adolescents' school-subject preferences over the life course. We find evidence for two mechanisms related to the effect of peers on favourite subjects: individuals' favourite subject preferences are affected both by the preferences of their friends (*friend influence effect*) and the preferences of their peers in the classroom (*peer exposure effect*). In this way, we managed to capture two substantively different ways in which the social surroundings of adolescents have an impact on individual subject preferences, while taking into account potentially important exogenous factors, such as parental education and cognitive ability. Both types of peer effects represent mechanisms that contribute to our understanding as to why more women drop out of a STEM career trajectory than men. High school is an important formative time, and preferences that are cemented

during this time are likely to be followed up on during tertiary education and further career steps.

Both boys and girls are influenced by their friends regarding their favourite subjects. The extent of this influence varies: Girls are influenced by their friends to a lesser extent than boys. Notably, this applies to all subjects, not just STEM subjects. To grasp the full significance of this friend influence that we document in our analysis, it has to be interpreted together with the observed gender homophily in the friendship networks, and with the already existing gender differences in subject preferences at the first observation.

First, as typical of adolescent friendship networks, friendship in our analysed sample is highly segregated by gender: As mentioned above, 90% of friendships are same-gender in wave 1, and 87% in wave 2. This entails that social influence on individuals mostly comes from friends of the same gender, i.e., girls will mostly be influenced by girls, and boys will mostly be influenced by boys. Second, in the first wave of our data, we already see a discrepancy in subject preferences between boys and girls: 19% of girls and 21% of boys had a STEM subject as a favourite subject; in wave 2, 20% of the boys still have a favourite STEM subject (5% decrease) but only 15% of the girls (21% decrease). Thus, boys are more likely to have a favourite subject that is a STEM subject in comparison to girls. This means that general social influence on favourite subject among boys will more likely be towards STEM subjects, and this influence affects mostly boys. While there is also influence on favourite subjects among girls, the fact that they are less likely to have a STEM subject as their favourite will lead to a comparatively lower influence towards STEM. Additionally, we find that social influence on favourite subjects is stronger among boys, which contributes to social influence amplifying gender differences in STEM preferences. Consequently, friend influence does not only produce gender differences in subject preferences because it works differently for boys and girls, but because a similar process reinforces different patterns for the two genders.

While we identify a general tendency that girls tend to like subjects that other girls also like, this relationship is a lot stronger for STEM subjects. This implies that the salience of gender norms in the classroom affects girls in their subject preferences. The

more girls are exhibiting non gender-normative preferences in a class, the lower are the costs of and the risk for sanctioning. Moreover, the effect of having other girls in the class who like STEM is potentially even stronger than it first seems, given that in these cases girls should, on average, also be more likely to have friends who like STEM. Therefore, in these cases friends could even have a more encouraging effect on girls' STEM-preferences. Peer exposure is thus a social mechanism that can help protecting girls from sanctions against non gender-conform behaviour.

Our study has several limitations. First of all, while we aimed at controlling for the most important alternative explanations, there are a few we were not able to take into account. For instance, another reason for the observed peer exposure effect could be a teacher effect: A science teacher that can encourage many girls in the classroom to engage and like the subject, is likely able to motivate additional girls as well. Our data, however, do not contain information about specific subject teachers along with subject preferences, only about the main class teacher; we were thus unable to account for this possibility. While we cannot make causal claims, both our interpretation of our results and the above alternative explanation suggest that an encouraging atmosphere for girls in the classroom regarding STEM subjects is important, and that is likely to influence female students. This could lead to girls not being discouraged from following their initial interest in STEM subjects.

Second, we are only focusing on Sweden, for technical and conceptual reasons: data quality and a non-tracked education system. We nevertheless believe that our results are likely to represent relevant tendencies in other countries, too. While Sweden is a country that has a relatively small gender gap in the proportion of science and technology jobs, we still find evidence for two social processes that lead to gender segregation in STEM preferences. Especially regarding the friend influence effect, our analysis considers all subjects, coupling this with gender homophily in the friendship network and initial STEM preferences. While gender homophily can be expected in virtually all adolescent friendship networks, gender differences in STEM preferences are likely to be more pronounced in other countries in comparison to Sweden. Indeed, this is the case in the other three countries in our data, i.e. England, Germany, and the Netherlands

(see Table C.5 in the Appendix). There is no reason to assume that the mechanisms exacerbating these differences, that we find evidence for in Sweden, are working differently in other countries. Nevertheless, more research is needed to investigate other national contexts beyond Sweden, and beyond other countries in Europe. Furthermore, it would be valuable to replicate this study for different age groups, to have a more complete view on the leaky pipeline, which assumes a life course approach.

Finally, we could not test whether the friend influence effect varies by subject and thus operates differently for STEM. Similarly, influences from friends of a specific gender could not be tested separately. This is because necessary effects for these tests are to date not implemented for the joint analysis of one-mode and two-mode networks in SAOMs. This, however, is necessary to capture interdependencies of specific subject preferences. Nevertheless, both notions would be interesting alleys to follow up on our results here, once these tests are possible.

Our results emphasise the effect of peers on individual subject preferences. While there is little that policy-makers can do about the friend influence that we document here, although awareness is valuable nevertheless, evidence for the peer exposure effect has potential in this regard. As mentioned above, the peer exposure effect suggests that it is beneficial for girls to be exposed to other girls that are interested in STEM subjects, potentially due to less or at least less costly sanctioning. Our analysis is not able to identify a threshold, such as a minimum number of girls that like STEM in a classroom that is needed to sufficiently lower associated costs. Future research focusing more specifically on this could provide insights into beneficial class compositions.

4.6 Conclusion

Through the application of the new and powerful method of multilevel longitudinal network analysis, our study demonstrates the importance of peers in explaining why women drop out from STEM careers at a higher rate than men. We identify two social processes that simultaneously contribute to the social reality of youths in the classroom, in which subject preferences of both girls and boys are shaped.

First, both boys and girls are influenced to like what their friends like. Since

students mostly have same-gender friends, on average, gender-specific tendencies of influence will emerge. Boys, who have higher probabilities to like STEM to begin with (in our observed time period), are likely to be further influenced towards STEM, because their friends are likely to be boys as well, who are, again, more likely to have STEM preferences themselves to begin with. Girls, having lower probabilities to like STEM already, are likely to be further influenced by other girls, who are also less likely to prefer STEM. However, these are just general tendencies: Depending on the particular friends one might have, individual implications can also be different, for example in cases where girls are friends with more boys than with girls (or with girls who like STEM).

Second, regarding STEM subjects, other girls' preferences in the classroom matter: having other female students in a class who prefer STEM can protect girls from getting discouraged from STEM subjects. This implies that besides subject preferences of friends, the negotiation of gender politics in the classroom is also important for girls' STEM preferences. Our findings suggest that the STEM pipeline model should be conceived as a social pipeline model, in which effects of peer exposure and friend influence are considered important factors of female dropout from STEM careers.

Chapter 5

Social influence on maths competence beliefs

While girls and boys perform similarly well in maths, girls have on average lower confidence in their maths skills, i.e. a lower perceived competence. This can result in lower motivation to engage with the subject, and in the choice to focus on other subject areas. Such differences in competence beliefs may in the long run play a role in occupational choices that lead to the large-scale occupational sex segregation in society, that is, to the tendency for men and women to work in different occupational sectors. Despite their importance and the attention that they have received in psychological theories we know very little about the social factors that shape differences in competence beliefs between boys and girls. Through the utilisation of multi-level stochastic actor-oriented models (SAOMs), this study analyses large-scale friendship network data from the Swedish subsample of the *CILS4EU* dataset (237 classrooms, 5,251 individuals, and 28,501 friendship ties at two time points) to analyse peer influence on maths competence beliefs of individuals, and the extent to which social dynamics apply differentially for boys and girls. Results show clear evidence for influence from friends on perceived performance in maths for boys and girls. Furthermore, boys are more likely than girls to be influenced by their friends to increase their maths competence beliefs, while objective measures (i.e. grades) of their maths skills are more influential for girls. This study thus shows specific social mechanisms that contribute to the gender gap in

maths competence beliefs.

5.1 Introduction

More girls than boys consider themselves to have very low maths skills, and more boys than girls consider themselves to have very high maths skills (Correll, 2004; Crombie et al., 2005). The vast gender gap is not mirrored by actual performances, since boys and girls tend to perform almost equally well in maths (e.g. Hyde, 2005; Sikora and Pokropek, 2012). Self-confidence in subjects informs educational choices, since it is reasonable to build one's career in a subject area that one considers oneself to be particularly good at. Hence, having high confidence in one's maths skills will render it more likely to build a career in a field that requires a sound mathematical background, such as the field of Science, Technology, Engineering, and Maths (STEM). Indeed, women tend to be underrepresented in this occupational field. Jobs in this area are considered to be demanding and cognitively challenging, and are remunerated comparably well. Thus, the fact that fewer women work in STEM fields has implications for the gender pay gap as well as for the general social status of women. This has received a lot of academic and political attention, motivated by a social equality perspective (since jobs in this field have a higher associated pay) and by a company-performance perspective (since the STEM occupational field loses out on potentially valuable employees). Understanding how gender differences in perceived maths skills come about is thus a central aspect of explaining occupational segregation.

This study investigates the social mechanisms that shape maths competence beliefs among adolescents. In particular, direct social influence from friends is investigated, i.e. whether a student changes his/her own competence beliefs so that they are in line with their friends' competence beliefs. Since traditional gender norms assume differences in suitability for the study and application of maths by gender, i.e. that girls are less suited, this study proposes and tests whether social influence is stronger when it is in line with traditional gender roles: Are girls more influenced by their friends in decreasing their maths confidence, and boys more in increasing their maths confidence?

Confidence in one's skills in a particular subject has received much attention from

empirical research in social psychology in the research of academic self-concepts. The theoretical model considers the embeddedness of how individuals evaluate their skills in particular areas with other facets of identity, which includes the dynamic interaction with significant others that surround the individual. Past research has tended to opt for the utilisation of classroom-level aggregates to define reference points for comparison and interaction, not differentiating between friends and others classmates, thereby being unable to identify significant others. This study is improving the fit of the theoretical and the empirical approach, by explicitly considering individuals and their friends by means of multilevel longitudinal social network analysis. The social dimension that the individual is embedded in and interacts with is captured by friendship network data, and the longitudinal feature of the data makes it possible to test the influence from friends. Through simultaneously analysing a large number of classroom's friendship networks with the newly developed multilevel framework for longitudinal social network models, this study is testing a particular social mechanism that has been implicitly assumed, but not yet explicitly tested, and investigates whether this applies differently to boys and girls. The main finding of the study is that social influence from friends on individually perceived maths skills applies differently to boys and girls in a way that contributes to the stabilisation of gender differences.

5.2 Theory, previous research, and hypotheses

This section gives an overview of relevant theory and previous research, and develops the hypotheses guiding the research presented in this chapter.

5.2.1 Maths competence beliefs

Beliefs regarding one's own competences and skills, such as in maths, has been researched within the framework of academic self-concept (ASC), and has been considered to belong to a number of self-concepts which are embedded in one's identity. While this chapter is not concerned with academic self-concepts in maths but only with maths competence beliefs, the theoretical considerations brought forward by previous research are insightful for competence beliefs as well. Academic self-concepts are

mental representations of one's skills, they are not just descriptive but evaluative, and make up the mental image of one's own person (Baumeister, 2011). This concept goes back to Shavelson et al. (1976) who proposed a general self-concept that consists of a variety of lower-order self-concepts. These facets of the general concept that a human being holds about him- or herself, are created through experiences and subsequent summarisation and categorisations by the individual. The four main categories cover individuals' perspectives regarding their i) social relations; ii) their emotional self, such as their self-esteem; iii) their physical self, such as their physical abilities and appearances; and iv) their academic self, i.e. academic skills, knowledge, and aptitude in a variety of different subject areas. The concepts on higher levels are considered to be rather stable, while single facets are more changeable since they are more specific. This results in individuals having different self-concepts for different subjects. The multidimensionality of the academic self-concept has received much empirical attention and supporting evidence (e.g. Marsh and Shavelson, 1985; Arens et al., 2011), and it has been found to affect a variety of outcomes.

5.2.2 Why this is important

Maths competence beliefs have been found to differ vastly by gender: Boys tend to have higher competence beliefs in maths and the sciences (Correll, 2004; Crombie et al., 2005), whereas girls have higher confidence in their reading skills, in languages, and social activities (Chow and Salmela-Aro, 2011). These differences in competence beliefs in high school have been repeatedly found to impact educational and career choices over the life course (Wigfield and Eccles, 2000; Watt et al., 2012), e.g. when it comes to continuation rates from secondary to tertiary education or to focussing on a particular subject, either regarding advanced courses in high school or study choices at university (Marsh and Yeung, 1997; Köller et al., 2006). Lower maths competence beliefs among girls thus lead women to avoid fields that are especially related to maths, such as jobs in STEM areas, which contributes to occupational gender segregation. This, in turn, tends to sustain the gender pay gap, since jobs in this field tend to be higher paying (Charles and Bradley, 2002).

Since maths competence beliefs are central to educational choices that lead to a STEM career, understanding how these differences come about is vital, probably more so than understanding the roots of gender gaps in actual performance in maths, because they are a lot smaller than the gender gap in maths competence beliefs (e.g. Hyde, 2005; Sikora and Pokropek, 2012). The fact that the gender gap in performance tends to appear and grow over time suggests that there are other mechanisms than cognitive ability that are contributing to this as well: Other factors, such as competence beliefs, are likely to contribute to this dynamic (Spelke, 2005).

5.2.3 What shapes perceived skills?

Competence beliefs are shaped by objective feedback on skills, and the way this feedback is judged by the individual. First, the higher the grades are that an individual gets, the higher the individual's confidence in his or her skills in that subject. Simultaneously, competence beliefs have been found to affect performance through higher motivational levels and trust in one's own skills (Huang, 2011; Marsh and Martin, 2011); this applies especially to subjects that have the reputation of being difficult, such as maths or the sciences (Wang and Degol, 2013). This reciprocal relationship has been documented repeatedly (e.g. Marsh and Craven, 2006; Möller et al., 2011; Niepel et al., 2014), and by itself already leads to an amplification of gender differences in competence beliefs. Since boys tend to perceive their maths competence higher than girls do, net of performance, boys will on average be more motivated to work and study for maths, and will increase or at least stabilise their performance. Their maths competence belief will thus increase further or remain more stable over time. Consequently, the reciprocal relationship between the competence belief and performance is likely to have an increasing or stabilising effect on the maths competence beliefs for boys.

Second, there several factors affecting the way the individual judges his or her skills in a specific subject: Individuals evaluate their skills in e.g. maths in comparison to the maths skills of others around them, but also in comparison to their skills in other subjects. Girls get higher grades in languages than they do in maths – they are, comparatively, worse in maths than they are at languages. Even though boys are not

substantially better than girls in maths, they get lower grades in languages (Jonsson, 1999). These tendencies likely contribute to the gender differences in perceived maths skills which are much larger than differences in performance.

These vast gender differences in maths competence beliefs are also likely to be informed by gender roles and gender identity. Competence beliefs, as understood by the academic self-concept, is embedded in a hierarchical, multidimensional structure that makes up the general self-concept. This embeddedness leads to facets interacting and affecting each other. Social Identity Theory, for instance, proposes that group membership or non-group membership, and the related social self-concept, affects individual identity and the perception of one's own skills (Hornsey, 2008). The gender identity of the individual, i.e. belonging to and identifying with the group of males or females, is likely to affect maths competence beliefs. Gender identity is informed by individual gender ideology, i.e. the "attitudes regarding the appropriate roles, rights, and responsibilities of women and men in society" (Kroska, 2007). Traditional gender roles suggest that boys are better at maths and analytical tasks, while girls are better at languages and are more communicative and nurturing (Jacobs et al., 2002). It has been shown that adolescents include these gender roles in their identity (Akerlof and Kranton, 2000; Sinclair and Carlsson, 2013). Since gender identity is embedded in the general self-concept, it can influence competence beliefs (Wigfield and Eccles, 2000). A recent study by van der Vleuten et al. (2016) finds evidence for individual gender ideology shaping girl's academic competence beliefs.

5.2.4 The influence of friends on maths competence beliefs

Friends have been repeatedly identified as the most important point of reference for adolescents (Youniss and Smollar, 1985), and the need to belong in the school community has been found to be integral to their well-being (Osterman, 2000). The importance of friends for individual outcomes has been found as such (Veenstra et al., 2013), as well as in relation to the entire classroom (Lomi et al., 2011).

While individuals tend to select friends based on similarity, such as the same gender, there has also been a lot of research documenting social influence, i.e. the tendency

of friend to become more alike over time. Several social mechanisms to explain why youths tend to conform to their friend's behaviour, their attitudes, and their beliefs. For example, there could be expected rewards due to conformity (Burgess and Akers, 1966), or in response to peer pressure (Brown et al., 1986), with individuals aiming at higher status or higher rates of acceptance from their friends. This can happen consciously or subconsciously through copying what individuals are exposed to, as proposed by Social Learning Theory (Bandura, 1977): If copied behaviour is rewarded in the particular social context, it is more likely to persist.

Friends are likely to discuss how interesting and how important maths is to them by itself and in relation to other subjects, including how skilled they view themselves. Since adolescents generally aim at fitting in with their friends, it is likely that similar dynamics also apply to maths competence beliefs. Thus, the first hypothesis that this research is testing is the following:

Hypothesis 1: Individuals adjust their maths competence beliefs towards that of their friends.

Since the interest of this paper is how social dynamics in this regard vary by gender, it will also be tested to what extent the hypothesised friend influence effect varies by gender.

At the same time, it is likely that there are gender differences in the direction of social influence that amplify this tendency even more. As discussed above, being good at maths is more compatible with the male gender role than with the female one. Research on the reinforcement of gender norms has found that individuals tend to remember and to be more perceptive to instances and attitudes that are in line with their own identity (Correll, 2001). This selective memory or selective perception is likely to interplay with the way maths competences are evaluated between friends, especially, since those friends are likely to be of the same gender, too: Social influence should thus be higher when it is in line with gender norms. Or, in other words, when friends are showing more gender-typed maths competence beliefs, social influence should be more effective.

This is in line with past research that has documented the tendency of educational choices to correlate with those of the same gender, but not with those of the opposite gender (Dryler, 1999). Thus, the second hypothesis that is tested in this paper is:

Hypotheses 2: The social influence of friends on maths competence beliefs is stronger when it is in the direction assumed by gender norms.

In other words, it is expected that social influence of friends is more influential for the maths competence beliefs of boys when it is increasing their beliefs, and vice versa for girls.

This study is specifically considering friends as the reference point for the social dimension of maths competence beliefs, which is more in line with theoretical considerations on the issue, and has not been investigated before. This is made possible through the availability of large-scale network data, as well as the newly developed methodological tool of random-coefficient, longitudinal social network models (Snijders and Koskinen, 2017). This method enables the separation of selection of influence effects, which is necessary when studying social influence from friends: Homophily in the variable of interest can come about either through friend selection (i.e. the tendency to befriend others that are similar to oneself) or through friend influence (i.e. the tendency to become more similar to one's friends). Thus, research aiming at identifying friend influence needs to control for friend selection, which would otherwise be confounded and lead to an overestimation of actual influence. This has already been possible for a while with Stochastic Actor-Oriented Models (SAOMs) that apply longitudinal social network analysis (Steglich et al., 2010), however, due to their single-network focus, findings cannot be generalised to a wider population. Random-coefficient multilevel network models are able to simultaneously analyse a large number of groups and to account for the multi-level structure of class-room data. Thus, by precisely measuring the social dynamics of maths competence beliefs for a large number of network groups, this paper is contributing to the literature on academic competence beliefs as well as to the explanation of gender differences in perceived maths competence, which has important

implications for macro-level patterns of horizontal occupational segregation.

5.3 Data and Methods

This section introduces the data and methods used in the analysis.

5.3.1 Data

The data used in this piece of research stems from the first two waves of the Children of Immigrants Longitudinal Survey in Four European Countries (*CILS4EU* Kalter et al., 2014, 2015), which contains data from England, Germany, the Netherlands, and Sweden. Due to technical reasons related to the suitability of the collected network data for the use in longitudinal social network analysis, only the Swedish part of the data is used in this study. The limits in usability are due to inconsistencies in nomination rules in England (Kruse and Jacob, 2014), and a restructuring of classrooms in the Netherlands, both leading to too much uncertainty and too little continuity in the classroom friendship networks. The German part of the data contains too many missing cases in either the friendship networks or other variables relevant to the analysis: Only about two-thirds of the classrooms would have been usable for the analysis; thus, this study focuses on Sweden only. Sweden is considered to be one of the most gender-egalitarian countries in Europe (Duvander and Andersson, 2006; Inglehart, 2008), which could have an effect on the size of the gender gap in in maths competence beliefs.

The Swedish part of the data contains information on 5,025 students from 251 classes in 129 schools (Kalter et al., 2014, 2015), aged 14/15 in wave 1, and 15/16 in wave 2. Of the 251 classes available, 10 were dropped from the analysis, for the following reasons: classrooms with fewer than 10 people, more than 70% missing values on maths competence beliefs in either wave 1 or 2, or no change in the friendship network. Two additional groups returned an error in the estimation. This leaves a total of 237 groups that are included in the analysis. In one class, there was too little variance in maths competence beliefs: All classroom members either had a value of 4 or missing information. In the estimation, missing values are internally replaced with the classroom mean, which was also 4 in this case, and which resulted in too little variance.

Table 5.1: Descriptive statistics on individual attributes, analytical sample.

	mean	SD	median	min	max	N
Maths competence beliefs, wave 1, low to high	3.74	0.97	4.00	1.00	5.00	4815
Maths competence beliefs, wave 2, low to high	3.64	1.04	4.00	1.00	5.00	4396
Grades in maths, low to high	2.49	0.81	2.00	1.00	4.00	4368
Gender: female						2647
Gender: male						2583

To include the class in the analysis, missing values were thus manually replaced by normally distributed values, making the information in the friendship network usable for the analysis, without distorting estimations in the influence part.

5.3.2 Variables

Descriptive statistics on these variables on individual attributes in the analytical sample can be found in Table 5.1, Table D.1 in the Appendix also presents these statistics of the full sample. There are no substantive differences between the distributions of the variables between the two samples.

The outcome of interest in this study are maths competence beliefs. This measure is based on the survey item "How well do you think you are doing in maths?", which was asked both wave 1 and 2. Answers were possible in five categories: "not well at all", "not that well", "okay", "quite well", and "very well".

Gender (male or female) is based on self-reports. The analysis furthermore controls for the grades in maths, which are based on register-data in Sweden and are only available at wave 2: They are given in four categories: "pass with high distinction" (MVG)", "pass with distinction" (VG), "pass" (G), and "fail" (U).

Friendship networks were constructed for analysis based on the survey item "Who are your best friends in class?". Five nominations were possible. Table 5.2 presents network level descriptives of the 237 networks that are included in the analysis; Table D.2 in the Appendix presents the same statistics for the classrooms that were excluded from the analysis, due to the reasons explained above. The excluded groups are somewhat smaller, since all classrooms with fewer than 10 people are excluded. The remainder

Table 5.2: Friendship network descriptives.

	min	max	mean	SD
size	10.00	34.00	22.11	4.32
density	0.01	0.43	0.13	0.04
degrees	0.00	5.00	2.71	1.85
average degree	0.12	4.70	2.66	0.76
clustering	0.00	1.00	0.57	0.12
jaccard index	0.04	0.58	0.32	0.10
same gender ties, wave 1	0.54	1.00	0.90	0.08
same gender ties, wave 2	0.51	1.00	0.87	0.10

of the descriptives does not give reason to assume substantive differences between the included and dropped groups.

Presented are the range, the mean, and the standard deviations of a number of statistics. On average, there are 22.11 individuals in a classroom. The average network density of 0.13 indicates that individuals are friends with on average 13% of their classmates. The average clustering coefficient of 0.57 indicates that 57% of two-paths are closed; i.e. 57% of two individuals that are friends with each other share another friend. The Jaccard coefficient is a measure for network change between the two waves: A coefficient of one indicates no change in the network, whereas zero indicates a complete change. The Jaccard coefficients in the friendship networks range from 0.04 to 0.58; 0.04 is very low, i.e. a very high change between wave 1 and 2. Again, due to the simultaneous estimation procedure, classrooms with low Jaccard coefficients can be included in the analysis. The friendship networks also show a strong tendency for gender homophily in both waves, with 90% of all friendship ties being between individuals of the same gender in wave 1, and 87% in wave 2.

5.3.3 Method and analytical strategy

This study deploys longitudinal social network analysis, which enables the separation between friend selection (friendship based on similar maths competence beliefs) and friend influence (similar maths competence beliefs based on friendship) processes. Furthermore, it allows to control for endogenous friendship network processes, such as becoming friends based on sharing another friend (transitive closure) or becoming friends

based on general popularity. Since these processes are interdependent, they violate independency assumptions of regression analyses and need to be taken into account in a network analysis framework.

The study uses Stochastic Actor-Oriented Models (SAOMs, (Snijders, 2011)). These models use simulations to estimate the tendencies that drive tie changes in the networks over time. It is simulated how actors change or keep their network ties based on these tendencies; based on this, social mechanisms can be inferred. SAOMs can be specified in a way that they do not only model how the network changes, but simultaneously in what way the behaviour of actors changes, and in what way network and behaviour interact (Steglich et al., 2010). In this way, friendship selection and influence processes can be separately accounted for. This analysis takes on this strategy and models the co-evolution of friendship networks and the maths competence beliefs.

In order to account for the multilevel structure of classroom-based educational data, the analysis is using newly developed random-coefficient multilevel framework for SAOMs (Snijders and Koskinen, 2017; Ripley et al., 2018). Analogous to the non-network hierarchical linear model for multilevel analysis, this approach additionally allows the analysis of the network-dependencies. Details on the estimation procedure can be found in section C in the Appendix.

All analyses have been carried out in *R*, version 3.2.2, using the package *RSienaTest*, version 1.2-2 on a large-scale computer cluster.

5.3.4 Model specifications

This study aims to test social influence of friends on perceived maths skills, and thus features a network (selection) part that models the friendship network, as well as a behaviour (influence) part that models math competence beliefs. In Model 1, to test social influence of friends on individual maths competence beliefs, an effect is included that models the tendency for the individual to change his or her maths competence belief in such a way that it becomes closer to the average of his or her friends (*friend influence*)[1]. While there are also other ways in which social influence on perceived

[1]Names in italics are used in the results tables. A full overview of all effects included in the analysis, as well as the internal *RSiena* shortnames, is given in Tables D.3 (selection part of the model) and D.4

maths skills could work, this conceptualisation of convergence is in line with the theoretical discussions on conformity among friends or friendship groups.To test whether this influence is different for boys and girls, an interaction with the gender variable, *effect from female*, is included (Model 2).

The second aim of this analysis is to test whether social influence on maths competence belief is directionally different for boys and girls. If not otherwise specified, SAOMs assume symmetrical effects, i.e. that parameters guide both behavioural increases and decreases. To test for differences in social influence on maths competence belief for boys and girls, a "creation" and an "endowment" version of the friend influence effect is used in a separate model, instead of the symmetrical friend influence effect. These are interacted with the gender variable *effect from female* (Model 3). In this model specification, the creation version of the friend influence effect, *friend influence (increase)* estimates the influence from peers on increasing competence beliefs, versus decreasing or not changing it. The endowment version, *friend influence (decrease)*, estimates the influence from peers on decreasing the level of competence beliefs, versus increasing or not changing it. Since the interest is whether this asymmetric influence differs by gender, *friend influence (increase)* and *friend influence (decrease)* are interacted with the gender variable, *effect from female*. A positive coefficient of the main effect *friend influence (increase)* and a non-significant coefficient of the main effect *friend influence (decrease)* would indicate that social influence for boys predominantly leads to an increase in perceived maths skills. A negative coefficient of the linear combination of the main effect *friend influence (decrease)* and its interaction with *effect from female* and a non-significant coefficient of the linear combination of the main effect *friend influence (increase)* and its interaction with *effect from female* would indicate that social influence for girls predominantly leads to a decrease in perceived maths skills. For a similar empirical analysis, see Haas and Schaefer (2014).

Furthermore, a number of controls is added to Models 1 to 3. First, to control for friendship selection, an effect that captures the tendency for individuals to become friends with others who have similar maths competence beliefs (f*riendship selection*) is

(influence part of the model) in the Appendix

included in the friendship part, along with the following structural: *reciprocity*, *transitive triplets*, *transitive reciprocated triplets*, degree-related effects, as well as effects to capture gender homophily. In the behaviour part, all models control for a general increase of maths competence beliefs and the effect of perceived maths skills on itself (*linear shape* and *quadratic shape*, respectively). Additionally, maths grades are controlled for as actor attributes in the behaviour part (*effect from maths grades*). Since it is assumed that they inform maths competence beliefs to a large extent, they are added to Models 1 to 3 in a step-wise manner, resulting in Models 1a, 2a, and 3a with all effects but maths grades, and Models 1b, 2b, and 3b with all effects including maths grades.

Based on visual inspection of the estimation trace plots, all models presented here show good convergence.

5.4 Results

Before the results from the multilevel network models are discussed, a few descriptive results based on the main variables of interest will be presented.

5.4.1 Descriptive Results

Tables 5.3 and 5.4 show cross-tabulations of maths grades and perceived maths skills in maths by gender, respectively. The patterns are in line with previous findings: While girls and boys do not achieve vastly different maths grades, boys have higher beliefs in their maths competences. It is the case that among those who receive the lowest maths grades there are more girls, however, the differences are small (52%); at the same time there are also more girls among those who receive the best grades in maths (51%). If competence beliefs were exclusively based on actual performance, no substantial gender gap should be expected. On the contrary though, more boys report a higher belief in their maths competence in comparison to girls, and that tendency increases: in wave 1, of those who have the highest maths competence beliefs, 56% are boys, and 58% in wave 2. Moreover, among those who report the lowest level of perceived maths skills, there are more girls than boys: In wave 1, 64% are girls, and in wave 2, 70% are

girls. These tendencies, i.e. absolute differences in the maths competence beliefs while showing very similar levels of performance, as well as the increase of the gender gap over time are in line with the presented theoretical considerations on endogenous and socially amplifying processes.

Figure 5.1 shows six exemplary friendship network plots from the pool of networks that are analysed in this study. Nodes are coloured according to individual maths competence beliefs, and shaped according to gender. The example plots serve as illustrations of the assumed tendencies: the clustering of friendship based on gender as well as similar perceived maths skills. Whether this impression applies, and whether the apparent correlation of friendship and perceived maths skills is due to social influence will be tested in the multilevel network analysis.

Figure 5.2 shows the change in maths competence beliefs between wave 1 and 2 by gender. Plotted is the difference between individual scores in the two waves, with negative values indicating a decrease from wave 1 to 2, and positive values an increase. Approximately half of the individuals change their maths competence beliefs between

Table 5.3: Maths grades (wave 2) by gender.

	boys	girls
Fail	0.08	0.08
Pass	0.48	0.47
Pass with distinction	0.32	0.32
Pass with high distinction	0.12	0.13

Table 5.4: Maths competence beliefs by gender.

Based on survey item: How well do you think you are doing in maths?

	wave 1		*wave 2*	
	boys	girls	boys	girls
not well at all	0.01	0.02	0.02	0.05
not that well	0.06	0.10	0.06	0.11
okay	0.27	0.28	0.30	0.31
quite well	0.40	0.39	0.35	0.34
very well	0.27	0.20	0.27	0.19

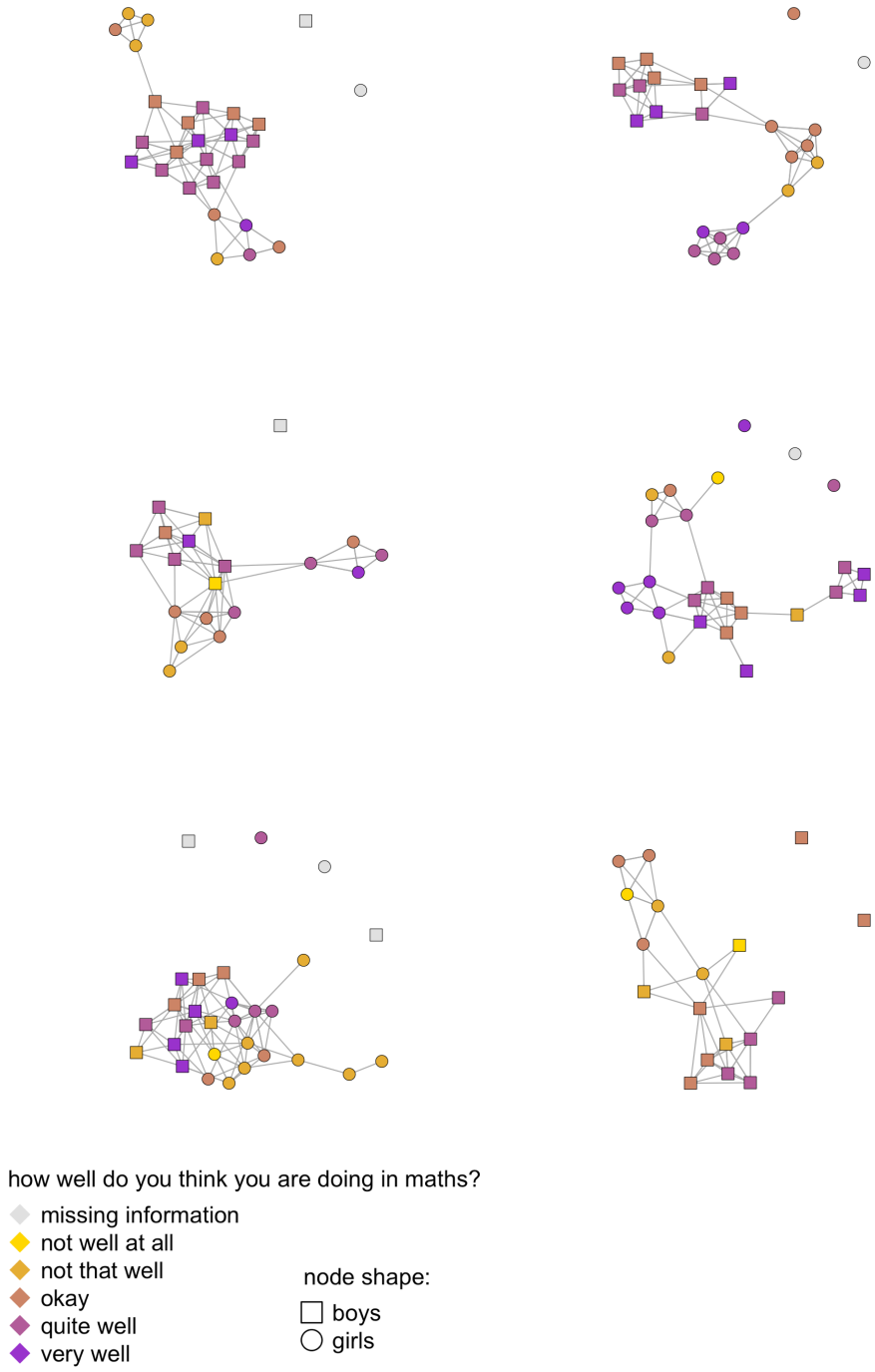

Figure 5.1: Six sample friendship network plots from the Swedish part of the *CILS4EU* data; nodes coloured according to maths competence beliefs.

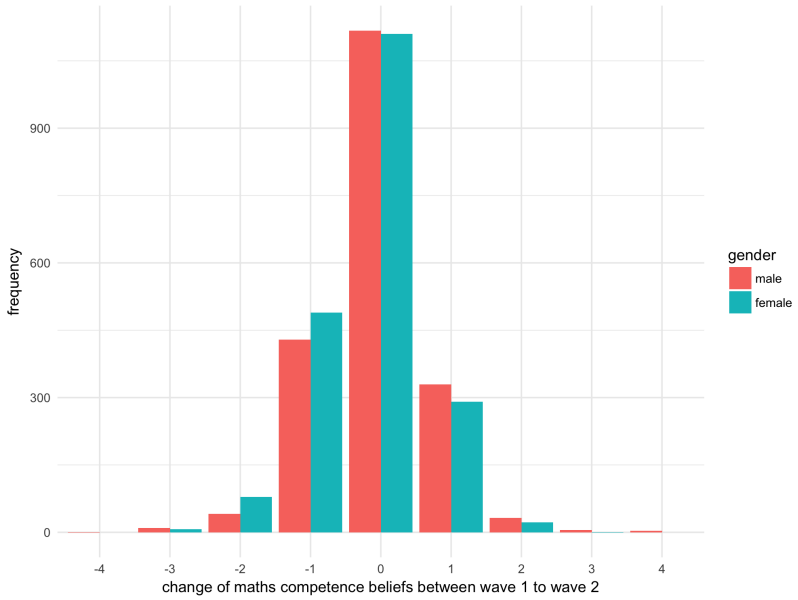

Figure 5.2: Change of perceived maths skills between wave 1 and 2, by gender.

the two waves, and most of them by one scale point. More girls decrease their competence beliefs, and more boys increase it, in line with the theoretical expectations.

5.4.2 Results from multilevel network models

Tables 5.5 to 5.7 show the results of interest; the full models can be found in Tables D.5 to D.10 in the Appendix. Generally, the results are in line with theoretical expectations: Girls tend to have lower maths competence beliefs, indicated by the negative coefficient of the effect from female variable in all models, including those that control for maths grades (Models 1b and 2b). Moreover, the models that control for maths grades show that maths grades are more important for girls in their effect on the perceived maths skills. In Model 1b, the effect of maths grades on maths competence beliefs is $\beta = 1.195$ for boys and $\beta = 1.195 + 0.236 = 1.431$ for girls; both effects are highly significant and significantly different from each other.

Hypothesis 1 expects that individuals adjust their maths competence beliefs towards those of their friends; this hypothesis is supported. *Friend influence* is positive and significant in Models 1a. In 1b, i.e. when controlling for maths grades, friend influence is still significant at the 10% significance level. Thus, there is evidence for the tendency for individuals to change in what way they perceive their maths skills towards the

Table 5.5: Results of Models 1a and 1b. Random coefficient multilevel stochastic actor-oriented models, friendship selection and math competence belief behaviour.

	Model 1a		**Model 1b**	
	β	SD	β	SD
Influence part				
Effect from female	-0.242***	(0.045)	-0.442***	(0.053)
Friend influence	0.854***	(0.228)	0.397^{+}	(0.296)
Effect from math grades			1.195***	(0.105)
Effect from maths grades * effect from female			0.236***	(0.083)

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{+}p < 0.1$

This model controls for friendship selection, endogenous network effects in the selection and the behaviour part. For full results see Table D.5 and D.6 in the Appendix.

Table 5.6: Results of Models 2a and 2b. Random coefficient multilevel stochastic actor-oriented models, friendship selection and math competence belief behaviour.

	Model 2a		**Model 2b**	
	β	SD	β	SD
Influence part				
Effect from female	-0.220***	(-0.045)	-0.436***	(-0.058)
Friend influence	1.635***	(-0.380)	0.508	(-0.492)
Friend influence * effect from female	-0.703*	(-0.293)	-0.053	(-0.386)
Effect from math grades			1.210***	(-0.114)
Effect from maths grades * effect from female			0.224***	(-0.089)

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{+}p < 0.1$

This model controls for friendship selection, endogenous network effects in the selection and the behaviour part. For full results see Table D.7 and D.8 in the Appendix.

Table 5.7: Results of Models 3a and 3b. Random coefficient multilevel stochastic actor-oriented models, friendship selection and math competence belief behaviour.

	Model 3a		**Model 3b**	
	β	SD	β	SD
Influence part				
Effect from female	-0.232***	(-0.054)	-0.437***	(-0.068)
Friend influence (decrease)	-0.081	(-0.627)	-0.369	(-0.762)
Friend influence (increase)	3.829***	(-0.789)	1.800*	-0.888
Friend influence (decrease) * effect from female	-0.416	(-0.533)	-0.213	(-0.635)
Friend influence (increase) * effect from female	-1.016+	(-0.651)	0.267	(-0.755)
Effect from math grades			1.198***	(-0.119)
Effect from math grades * effect from female			0.235*	(-0.101)

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{+}p < 0.1$

This model controls for friendship selection, endogenous network effects in the selection and the behaviour part. For full results see Table D.9 and D.10 in the Appendix.

average of their friends. The interaction of *friend influence* with *effect from female* in Model 2a indicates that friend influence effect varies significantly by gender: $\beta = 1.635$ for boys and $\beta = 1.635 + -0.703 = 0.932$ for girls. When controlling for maths grades in Model 2b, both the main and the interaction effect are no longer significant.

Hypothesis 2, stating that social influence of friends on maths competence beliefs is stronger when it is in the direction assumed by gender norms, is not supported by the results. In Model 3a, the main effect of *friend influence (increase)* is positive and highly significant, and its interaction with *effect from female* is negative and marginally significant. This indicates that *both* boys and girls tend to increase their maths competence beliefs towards the average of their friends, instead of remaining unchanged or decreasing towards the average of their friends, and that this tendency is weaker for girls: $\beta = 3.829$ for boys and $\beta = 2.813$ for girls. In line with the findings from Models 1 and 2, the tendency for girls to be more perceptive towards maths grades is stronger in comparison to boys.

Friend influence (decrease), however, is not significant as the main effect, nor is its

interaction with effect from female. This indicates that students are not influenced by their friends towards decreasing their competence beliefs in maths, and that this does not vary by gender. This shows that neither boys nor girls tend to decrease their maths competence beliefs in maths towards the average of their friends instead of increasing it or keeping in unchanged: When friends have on average lower maths competence beliefs, boys and girls are unlikely to lower their own competence beliefs so that they are closer to the ones of their friends.

5.5 Discussion

This study set out to investigate the social dynamics shaping maths competence beliefs, specifically whether friends influence each other, and whether this varies by gender. Descriptively, we could verify some previous results. For example, girls tend to evaluate their maths skills lower than boys, even when controlling for relatively objective measures of their skills, i.e. maths grades. As discussed in the literature review, this tendency by itself will already lead to an amplification of individual tendencies throughout the school career, due to the reciprocal relationship between perceived competence and performance: The higher the perceived competence, the higher the motivation to study, and thus the higher the performance and the higher the perceived competence. Descriptively, there are already gender differences in the patterns of perceived maths skills in the data analysed for this study, which grow between wave 1 and 2. Importantly, there are no substantial gender differences in the maths grades that could warrant this gender gap in perceived maths skills; this is suggestive of girls and boys differently appraising their skills, for example in relation to their skills in other subjects. The size of the gender gap, which is documented in the data here, is also likely due to the fact that the individuals that are analysed are adolescents: The mechanisms that are proposed here have thus already been adding to amplification in the past.

The biggest contribution in this study is the ability to identify social influence effects from friends, using SAOMs to control for selectivity in a large number of network groups. Studies in the field of longitudinal social network analysis (Steglich et al., 2010) have been able to identify particular social mechanisms due to their focus on

social ties themselves, but they usually had to rely on single or small numbers of social networks. Regression studies in mainstream educational sociology, on the other hand, are conducted on large-scale, representative, in part cross-national data. Some rely on classroom-level aggregates as proxies for their hypothesised social mechanisms, and some include friend characteristics in the regressions, neglecting the endogenous nature of friendship dynamics over time. Both approaches make valuable contributions. Longitudinal social network analysis can find strong evidence of particular social mechanisms of influence, which regression analyses cannot achieve; regressions can generalise their results to a wider population, which social network analyses with a narrower focus are unable to do.

In this study, both are achieved: the focus on particular social mechanisms, while analysing a large data set that allows for the generalisation of results. This is possible through a novel addition to the framework SAOMs, that now allows for the analysis of a large number of separate networks (e.g. school classes), jointly in a random coefficient multilevel model (Snijders and Koskinen, 2017; Boda, 2018). Compared to earlier solutions within the SAOM framework, that were typically based on the meta-analysis of group-level results, this approach offers much more statistical power. This is crucial, since school classes are normally rather small. A lack in statistical power has consequently been a major issue when analysing social influence in educational settings using SAOMs (Stadtfeld et al., 2018).

The two main findings are that friends do tend to influence perceived maths skills and that gender norms tend to not matter for social influence on perceived maths skills among girls.

This general tendency of social influence from friends on individually perceived maths skills will by itself amplify already existing gendered patterns. This is due to the fact that friendship networks in high school tend to be segregated by gender; this is also the case in the data, which is analysed here: In wave 1, 90% of all friendship ties are between individuals of the same gender in wave 1, and 87% in wave 2. Thus, girls are on average mostly influenced by girls, and boys by boys. Since social influence towards the average happens mostly within the same gender, this will lead to a

reduction in the variance of perceived maths skills among boys and among girls. If all social influence was exclusively happening within gender, and all else equal, this would lead to a bipolar distribution, with a non-overlap of the perceived skills of boys and girls. Although friendships are not completely segregated by gender, there is a strong tendency for gender homophily, which over time will still lead to the tendency of girls' and boys' perception of their respective skills to become more uniform. This can reinforce perceptions of gender roles, i.e. girls being less suited for maths than boys.

Regarding social influence, results are weaker when controlling for maths grades and when social influence is interacted with gender, which could be due to limited statistical power. Considering that the analysis does find statistically significant results for social influence when it is not interacted with gender and when it is split up by increase and decrease, the analysis overall does provide strong evidence for social influence on maths competence beliefs, which is found for both boys and girls.

In addition to this general mechanism, this study found that for girls, social influence is not more effective depending on whether it is in line with gender roles: Girls, as boys, tend to be influenced by their friends to increase their perceived skills. The fact that no evidence is found for a decreasing social influence effect for girls could be due to limited statistical power, but it is in line with previous research identifying gender roles to be less fluid for men than for women (England, 2010).

Descriptively, however, we see that boys on average increase, and girls on average decrease their perceived maths skills. While the increase among boys can be explained by the social mechanisms, the decrease among girls cannot. The analysis identified another tendency, which could further shed light on this: For girls, maths grades are more important. Coupling this finding with the above-described common tendency of gender homophily in friendship networks, this entails that boys will be influenced more towards increasing their maths competence beliefs towards the average of their friends, who likely will already have higher levels of maths competence beliefs than girls will. Girls are influenced by their friends to increase their maths competence beliefs as well, but are diverted more easily if they get worse grades. Since maths grades are less determinative for boys, this overall could contribute to the observed widening of

the gender gap in maths competence beliefs. Further research is needed to investigate these specific dynamics.

This study also knows some limitations. First, it was not possible to control for a number of factors that could further have shed light on the details of the social mechanisms proposed here. The model did not converge when effects controlling for performances and perceived skills in other subjects were additionally included. Second, grades in maths were only available at wave 2. This did not have severe implications, since the *RSiena* estimation routine does not allow for time-varying covariates when only analysing two waves, assuming that one observation, at time point one or 2, affects the change in the variable of interest. Third, the analysis was only carried out with data from one country. Lastly, one alternative mechanism that is likely at play in the described dynamic as well could not be addressed in this work: social comparison among friends. It is plausible to assume that if one's friends are very good at maths, one could feel comparably worse. However, studies at classroom level suggest that social comparison effects on educational choice are quite limited in scope (Jonsson and Mood, 2008). The analysis at hand is unable to account for such tendencies due to the mentioned convergence issues; the addition of new effects in *RSiena* could potentially address this, however, such alternative effects are not currently available.

Since this study identifies social mechanisms that stabilise existing differences, it can be suggested that similar dynamics apply in other countries since gender differences are likely to exist as well; this, however, should be tested by further research. This emphasises the importance of collecting more well-managed social network panels, so that future research can test the social mechanisms proposed and investigated here for other country-contexts as well.

5.6 Conclusion

This study investigates in detail some of the social mechanisms that are shaping adolescents' maths competence beliefs. Results show that social influence from friends and objective performance measures of maths skills can contribute to a widening of the gender gap in perceived maths skills: Social influence tends to lead to an increase in

maths competence beliefs for both boys and girls, but is stronger for boys. Furthermore, maths grades are more important for girls, and friend influence tends to happen mostly within the same gender, due to gender homophily in adolescent friendship networks. This results in a scenario in which boys are more likely than girls to be influenced by their friends to increase their math competence beliefs towards the average, whereas girls are more likely than boys to follow more objective measures of their skills. This also points to gender roles being less influential for girls, which is to their advantage. These findings show that social factors influence individual outcomes, which can have important implications for macro-level patterns of occupational segregation.

Chapter 6

Findings and discussion

This thesis proposes to specifically consider the social dimension of educational outcomes, and shows in four empirical chapters that social aspects are important for understanding educational success of the individual, and thus educational inequality on the level of society. Both the ego-centric as well as the socio-centric approach provide evidence for social interactions being related to performance in school, university aspirations, favourite subjects, and competence beliefs. Social aspects can furthermore contribute to the understanding of variations in these outcomes by gender and immigrant background.

6.1 Results and limitations

There are four main conclusions from this thesis. First, studying individual interactions, both in the ego-centric and socio-centric approach, is a valuable approach. Chapters 3 and 4 consider individual interactions as well as the wider classroom context (i.e. aggregates), and both find evidence for the importance of interactions with peers, for university aspirations and favourite subjects, respectively, net of the contextual classroom effect. This corroborates the value of past research that has documented these contextual effects, and the theoretical value of the Wisconsin Model that specifically proposes the influence of "significant" (and thus not all) others. While Chapter 3 shows clustering based on university aspirations within classrooms, Chapter 4 is able to separately account for friend influence (and selection) effects on favourite subjects and

for peer exposure, i.e. influence from all others in the classroom. The analysis finds evidence for both processes, net of each other.

Second, Chapters 2 and 3 show that individual interactions matter for both school performance and university aspirations, and that there is some evidence that points to these dynamics being different for youths with an immigrant background. Chapter 2 considers social isolation (in the friendship network) and social rejection (negative relationships), and finds a negative relationship with school performance for both natives and migrants, but that is stronger for newly arrived immigrants. Chapter 3 documents the general tendency of individuals to change their university aspirations in the same way their friends and peers do, and finds that this tendency applies less to immigrants. Even though Chapter 2 only finds a stronger effect of social interactions for newly arrived migrants and not for everyone with a migration background, the strong evidence of Chapter 3 gives reason to assume that social aspects are important factors in explaining differential educational success by migrants.

Third, Chapters 4 and 5 provide strong evidence for how social interactions can affect existing gender differences when it comes to favourite subjects and maths competence beliefs in maths. Both chapters show a general tendency for social influence from friends on educational outcomes, and both chapters document a tendency of this influence to be weaker for girls. These results have to be interpreted together with the fact that girls tend to be friends with girls, and boys tend to be friends with boys: In adolescence in general, and in the analyses in particular. This results in boys mainly influencing boys, and girls mainly influencing girls, albeit to a lesser extent. Regarding favourite subjects, this leads to an amplification of gender differences in individuals liking STEM subjects. Social influence and gender homophily as such, however, do not lead to such an amplification in maths competence beliefs, but to a stabilisation of gender differences. Evidence shows the tendency towards a convergence to the average of friends, which leads to a reduction in variance among boys and girls. This could strengthen perceptions of gender-typed behaviour. Chapter 5 additionally finds that while boys and girls both tend to increase confidence in their maths skills towards the average of their friends, girls are discouraged more easily than boys when they receive

lower grades.

Fourth, this thesis shows the benefit of using the socio-centric instead of the ego-centric approach when analysing social aspects of educational inequality: The analyses using the socio-centric approach find much clearer evidence for social mechanisms. This is in line with the theoretical considerations above, namely that social interactions and social structures are interdependent, and that both affect individual educational success. Since both socio-centric analyses consider gender, however, this could also mean that social aspects are more important when explaining gender differences in educational success instead of immigrant differences. However, the methodological limitations of especially Chapter 2 together with the greater complication of the migrant case due to a higher number of ethnic groups in comparison to gender, give reason to expect stronger results with a socio-centric approach. This, however, requires further research.

In the big picture, this thesis asked whether social interactions and social structures in the classroom can help explain differences in educational outcomes between boys and girls, and between children of natives and children of immigrants, respectively. These social aspects have often been theoretically utilised to explain educational inequality, but have received comparably little empirical attention, often due to the lack of suitable data. All four empirical chapters document distinct social mechanisms and their implications that are proposed in much theoretical work. Thus, it can be generally concluded that social interactions and social structures matter for individual educational success, as proposed by theory. Although these individual outcomes cannot be directly aggregated to the educational inequalities that are documented at the macro-level, differential individual success plays an important role in this dynamic; documenting how these individual outcomes are embedded in social structure and interactions is thus the main contribution that this thesis makes to the field.

Besides these achievements, there are some limitations, too. First, not the same methodological attention was spent on each aspect of the bigger question. Specifically, this refers to the fact that both chapters that focus on ethnic inequalities are using the ego-centric approach, and are thus not able to equally consider social structure

besides social interactions. This is mostly due to the nature of the multilevel network models that are utilised for the socio-centric approach: the analysis of gender, which is measured as a binary category in the data, is easier to handle by these models than ethnicity, that has more categories. This is especially true for Sweden, which is the country with the best network data quality in the *CILS4EU* data. Future studies, however, could address this limitation further; since there are constantly new opportunities realised by the developers of *RSiena*, the software used to apply stochastic actor-oriented models, it might be possible to fill this gap in the near future.

Furthermore, while this thesis is able to map concrete social mechanisms in a variety of country contexts, it is limited to four country contexts in Western Europe; in the case of the socio-metric approach, even only to one country. This is due to the fact that social network data is not usually part of large national and international surveys and highlights the need for further data collection.

Additionally, a limitation that came up in all the chapters is that there might be class-level effects, such as ethnic or gender composition, number of cluster or degree of segregation, that affect the social embeddedness of the individual in the classroom. Future research could shed light on this matter, which would be particularly interesting in the case of the multilevel network models.

6.2 Country differences in results

Both Chapter 2 and 3 identify some differences across countries in the results. Chapter 2 finds ethnic disadvantages that are mostly in line with the expectations. In particular, social aspects, here social exclusion, have the strongest effects where disadvantages are also strongest: Disadvantages accumulate. An unexpected pattern, however, is that the results for social exclusion are weakest in England. While there is no theoretical reason why this is the case, there is a practical one. In the different countries that are sampled in the *CILS4EU* data, the socio-metric survey units, i.e. exactly who could be nominated in the collection of the network data, is slightly different, and these differences are particularly pronounced in England. The biggest issue here is that in approximately a third of all 208 classrooms, it was not possible to nominate students

that were absent on that day, for example due to illness (Kruse and Jacob, 2014). This can lead to a substantial bias in measuring the friendship and avoidance network, which essentially leads to differences in the accuracy of measuring the social exclusion reality in each classroom. Since this analysis built the social exclusion measures on the number of ties, i.e. counts of nomination, it is likely that in the case of England, the extent of social exclusion is underestimated, and the results thus weaker. Dropping those classes was not desirable due to the large number of cases that would have been lost and associated issues with statistical power. An alternative explanation could be that there are comparably more different school types in England, which could render the cases not comparable and mask differences.

In Chapter 3, country-specific analyses find weaker results for clustering of university aspiration within the school classes for Germany and the Netherlands. As discussed above, this is likely the case because the education system in these two countries is tracked: Only students graduating from the highest tracks are typically eligible for university studies (although there are routes back to the academic track). Nevertheless, since the pooled analysis, using country-weights, finds overall evidence for clustering of aspirations within classrooms, it could also be due to lower statistical power.

As mentioned above, Chapters 4 and 5 only use the Swedish part of the data. While it is unfortunate that these analyses could not use data from all four countries, it was not possible to do so. The socio-metric approach as applied in these two chapters comes with a practical limitation: It requires well-managed social network data, and groups that do not change in their composition too much over time. Of the *CILS4EU* data, the Swedish data was the only one that met these requirements. England had issues with different nomination conditions in classes, and in the Netherlands, the composition change is too high due to differential sorting of classrooms that happens between the waves. Carrying out the analysis for Germany was attempted for Chapter 5, but due to many missings on crucial variables, including network variables for which imputation is not an option yet, more than a third of classes would have been needed to be excluded. For Chapter 4, the tracked education system in Germany returned too many different favourite subjects with partly very low numbers of cases, so that a uniform coding was

not realistic. This underlines the value of the ego-centric approach, i.e. the utilisation of regression analyses, which can handle missing data and do not have such strict data requirements. Moreover, the ego-centric approach can use survey weights in the analysis, to make claims about representativeness of the results.

Besides these identified cross-country differences, the main social mechanisms do still seem to apply in most instances, and, on a high level, only seem to vary in strength. This points to the identified effects being generic effects or associations in Western European countries.

Appendix A

Appendix Chapter 2

Table A.1: Ethnic groups by migration generation.

	native	newly arrived	first generation	second generation	total
United Kingdom	1994.00	2.00	8.00	5.00	2009.00
Germany	2211.00	7.00	15.00	24.00	2257.00
Netherlands	22.00	1.00	9.00	8.00	40.00
Sweden	2151.00	0.00	0.00	1.00	2152.00
Turkey	75.00	21.00	92.00	843.00	1031.00
Russian/Former USSR	63.00	37.00	187.00	136.00	423.00
Balkans	81.00	37.00	103.00	423.00	644.00
Southern/Eastern European	497.00	81.00	114.00	322.00	1014.00
Middle Eastern	50.00	90.00	230.00	473.00	843.00
Other Western/Northern European	551.00	9.00	26.00	124.00	710.00
Non-European Western	59.00	4.00	9.00	15.00	87.00
South Asian	121.00	41.00	119.00	504.00	785.00
Southeast/East Asian	73.00	53.00	90.00	120.00	336.00
Latin/South American	178.00	44.00	75.00	126.00	423.00
Africa	114.00	96.00	173.00	377.00	760.00
Oceania	1.00	0.00	1.00	1.00	3.00
total	8241.00	523.00	1251.00	3502.00	13517.00

Table A.2: Linear regression of grade point average (standardised). Non-imputed data.

	Model 1	Model 2	Model 3
migration background (ref: native)			
newly arrived	-0.41***	-0.26***	-0.21**
	(0.07)	(0.07)	(0.07)
1st generation	-0.19***	-0.14**	-0.12**
	(0.04)	(0.05)	(0.05)
2nd generation	-0.13***	-0.07	-0.06
	(0.03)	(0.04)	(0.04)
gender (female)		0.08***	0.09***
		(0.02)	(0.02)
age (in months)		-0.01***	-0.01***
		(0.00)	(0.00)
cognitive ability		0.07***	0.07***
		(0.00)	(0.00)
highest parental degree (ref: primary school)			
lower secondary school		0.09*	0.08
		(0.04)	(0.04)
upper secondary school		0.17***	0.16***
		(0.04)	(0.04)
university degree		0.52***	0.50***
		(0.04)	(0.04)
parental unemployment		-0.03	-0.02
		(0.02)	(0.02)
other language spoken at home		0.08*	0.09**
		(0.03)	(0.03)
school resistance		-0.06***	-0.06***
		(0.00)	(0.00)
isolated from everyone			-0.13
			(0.16)
isolated from majority students			-0.09**
			(0.04)
avoidance by majority			-0.32***
			(0.07)
avoidance by minority			-0.03
			(0.04)
number of friends (indegree)			0.02***
			(0.01)
intercept	0.17***	0.93**	0.97**
	(0.02)	(0.31)	(0.32)
adj. R^2	0.01	0.21	0.22
N	8376	8376	8376

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

All models control for survey country.

Table A.3: Linear regression of grade point average (standardised), with migration background and social exclusion interaction effects. Non-imputed data.

	Model 1	Model 2	Model 3	Model 4
migration background	-0.05	-0.05	-0.07	-0.05
	(0.04)	(0.04)	(0.04)	(0.04)
gender (female)	0.13***	0.13***	0.13***	0.13***
	(0.02)	(0.02)	(0.02)	(0.02)
age (in months)	-0.01***	-0.01***	-0.01***	-0.01***
	(0.00)	(0.00)	(0.00)	(0.00)
cognitive ability	0.07***	0.07***	0.07***	0.07***
	(0.00)	(0.00)	(0.00)	(0.00)
lower secondary school	0.11*	0.11*	0.11*	0.11*
	(0.04)	(0.04)	(0.04)	(0.04)
upper secondary school	0.17***	0.17***	0.17***	0.17***
	(0.04)	(0.04)	(0.04)	(0.04)
highest parental degree (ref: primary school)				
university degree	0.52***	0.52***	0.52***	0.52***
	(0.04)	(0.04)	(0.04)	(0.04)
parental unemployment	-0.03	-0.03	-0.03	-0.03
	(0.02)	(0.02)	(0.02)	(0.02)
other language spoken at home	0.09*	0.09*	0.09*	0.09*
	(0.04)	(0.04)	(0.04)	(0.04)
isolated from everyone	-0.05	-0.05	-0.05	-0.05
	(0.19)	(0.15)	(0.15)	(0.15)
isolated from majority students	-0.09*	-0.02	-0.09*	-0.09*
	(0.04)	(0.11)	(0.04)	(0.04)
avoidance by majority	-0.40***	-0.40***	-0.52***	-0.40***
	(0.07)	(0.07)	(0.10)	(0.07)
avoidance by minority	-0.01	-0.01	-0.01	-0.01
	(0.04)	(0.04)	(0.04)	(0.04)
number of friends (indegree)	0.02***	0.02***	0.02***	0.02***
	(0.01)	(0.01)	(0.01)	(0.01)
isolated from everyone * migr.b.	0.01			
	(0.33)			
isolated from majority students * migr.b.		-0.08		
		(0.11)		
avoidance from majority * migr.b.			0.22	
			(0.13)	
avoidance from minority * migr.b.				0.01
				(0.12)
intercept	0.71*	0.71*	0.73*	0.72*
	(0.32)	(0.32)	(0.32)	(0.32)
adj. R^2	0.19	0.19	0.19	0.19
N	8376	8376	8376	8376

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

All models control for survey country.

Abbreviations: migr.b. = migration background.

Table A.4: Linear regression of grade point average (standardised), England.

	Model 1	Model 2	Model 3
ethnic and migration background (ref: native)			
newly arrived	-0.26*	-0.36***	-0.33**
	(0.12)	(0.11)	(0.11)
1st gen non-western	-0.02	-0.11	-0.10
	(0.08)	(0.08)	(0.08)
1st gen western	-0.38	-0.36	-0.34
	(0.20)	(0.20)	(0.21)
2nd gen E.EU, RU, ME	0.16	0.15	0.13
	(0.21)	(0.14)	(0.14)
2nd gen Asian	0.03	-0.06	-0.05
	(0.06)	(0.08)	(0.08)
2nd gen African	0.27*	0.01	0.02
	(0.12)	(0.11)	(0.12)
2nd gen other	0.11	0.01	0.01
	(0.13)	(0.10)	(0.10)
gender (female)		0.00	0.01
		(0.04)	(0.04)
age (in months)		-0.00	-0.00
		(0.00)	(0.00)
cognitive ability		0.10***	0.10***
		(0.01)	(0.01)
lower secondary school		0.04	0.02
		(0.07)	(0.08)
other language spoken at home		0.20**	0.21**
		(0.07)	(0.07)
school resistance		-0.05***	-0.05***
		(0.01)	(0.01)
isolated from everyone			-0.09
			(0.22)
isolated from majority students			-0.06
			(0.06)
Avoidance by majority			-0.19
			(0.16)
Avoidance by minority			0.05
			(0.07)
number of friends (indegree)			0.02
			(0.01)
intercept	0.10***	-1.49*	-1.38
	(0.03)	(0.74)	(0.77)
adj. R^2	0.01	0.31	0.31
N	8376	8376	8376

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

All models controls for parental education.

Table A.5: Linear regression of grade point average (standardised), Germany.

	Model 1	Model 2	Model 3
ethnic and migration background (ref: native)			
newly arrived	-0.40**	-0.18	-0.10
	(0.14)	(0.14)	(0.14)
1st gen non-western	-0.31***	-0.20	-0.17
	(0.09)	(0.11)	(0.11)
1st gen western	-0.13	-0.14	-0.13
	(0.09)	(0.10)	(0.10)
2nd gen Turkish	-0.36***	-0.35***	-0.30***
	(0.05)	(0.08)	(0.08)
2nd gen Russian, Eastern EU	-0.16*	-0.18*	-0.18*
	(0.07)	(0.08)	(0.08)
2nd gen Balkans, Middle East	-0.18*	-0.10	-0.08
	(0.09)	(0.10)	(0.10)
2nd gen other	-0.04	-0.10	-0.08
	(0.11)	(0.12)	(0.11)
gender (female)		0.08*	0.08*
		(0.03)	(0.04)
age (in months)		-0.02***	-0.02***
		(0.00)	(0.00)
cognitive ability		0.05***	0.05***
		(0.00)	(0.00)
parental unemployment		0.02	0.04
		(0.04)	(0.04)
other language spoken at home		0.11	0.12
		(0.06)	(0.06)
school resistance		-0.06***	-0.06***
		(0.01)	(0.01)
isolated from everyone			-0.12
			(0.36)
isolated from majority students			-0.22***
			(0.06)
avoidance by majority			-0.26**
			(0.10)
avoidance by minority			-0.06
			(0.08)
number of friends (indegree)			0.02
			(0.01)
intercept	0.13***	2.89***	2.85***
	(0.02)	(0.43)	(0.43)
adj. R^2	0.02	0.15	0.15
N	8376	8376	8376

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

All models control for school track.

All models controls for parental education.

Table A.6: Linear regression of grade point average (standardised), Sweden.

	Model 1	Model 2	Model 3
ethnic and migration background (ref: native)			
newly arrived	-0.57***	-0.26*	-0.23
	(0.13)	(0.13)	(0.13)
1st gen non-western	-0.40***	-0.25**	-0.24*
	(0.08)	(0.09)	(0.09)
1st gen western	0.07	0.01	0.02
	(0.15)	(0.12)	(0.12)
2nd gen Turkish Middle East	-0.25***	-0.04	-0.04
	(0.07)	(0.08)	(0.09)
2nd gen RU, Balk, SE EU	-0.15*	0.03	0.03
	(0.07)	(0.08)	(0.09)
2nd gen Western	-0.21	-0.05	-0.05
	(0.12)	(0.10)	(0.10)
2nd gen other	0.09	0.08	0.06
	(0.12)	(0.12)	(0.13)
2nd gen African	0.00	0.23	0.22
	(0.11)	(0.12)	(0.13)
gender (female)		0.13***	0.14***
		(0.03)	(0.03)
age (in months)		-0.00	-0.00
		(0.00)	(0.00)
cognitive ability		0.07***	0.07***
		(0.00)	(0.00)
parental unemployment		-0.04	-0.04
		(0.05)	(0.05)
other language spoken at home		0.03	0.03
		(0.06)	(0.06)
school resistance		-0.07***	-0.07***
		(0.01)	(0.01)
isolated from everyone			-0.05
			(0.26)
isolated from majority students			-0.01
			(0.06)
avoidance by majority			-0.51***
			(0.13)
avoidance by minority			-0.11
			(0.07)
number of friends (indegree)			0.02
			(0.01)
intercept	0.21***	-0.49	-0.25
	(0.02)	(0.65)	(0.67)
adj. R^2	0.02	0.26	0.26
N	8376	8376	8376

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

All models controls for parental education.

Table A.7: Linear regression of grade point average (standardised), with migration background and social exclusion interaction effects, England.

	Model 1	Model 2	Model 3	Model 4
migration background (dummy)	-0.07	-0.06	-0.08	-0.06
	(0.06)	(0.06)	(0.07)	(0.07)
gender (female)	0.01	0.01	0.01	0.01
	(0.04)	(0.04)	(0.04)	(0.04)
age (in months)	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)
cognitive ability	0.10***	0.10***	0.10***	0.10***
	(0.01)	(0.01)	(0.01)	(0.01)
highest parental degree (ref: primary school)				
lower secondary school	0.03	0.03	0.03 v0.03	
	(0.08)	(0.08)	(0.08)	(0.08)
upper secondary school	0.20***	0.20***	0.20***	0.20***
	(0.06)	(0.06)	(0.06)	(0.06)
university degree	0.58***	0.58***	0.58***	0.58***
	(0.06)	(0.06)	(0.06)	(0.06)
parental unemployment	-0.07	-0.07	-0.07	-0.07
	(0.04)	(0.04)	(0.04)	(0.04)
other language spoken at home	0.20**	0.19**	0.20**	0.20**
	(0.06)	(0.06)	(0.06)	(0.06)
school resistance	-0.05***	-0.05***	-0.05***	-0.05***
	(0.01)	(0.01)	(0.01)	(0.01)
isolated from everyone	-0.07	-0.13	-0.13	-0.14
	(0.26)	(0.22)	(0.22)	(0.22)
isolated from majority students	-0.07	0.08	-0.07	-0.07
	(0.06)	(0.18)	(0.06)	(0.06)
avoidance by majority	-0.21	-0.20	-0.20	-0.20
	(0.16)	(0.16)	(0.22)	(0.17)
avoidance by minority	0.05	0.05	0.05	0.07
	(0.07)	(0.07)	(0.07)	(0.07)
number of friends (indegree)	0.02*	0.02*	0.02*	0.02*
	(0.01)	(0.01)	(0.01)	(0.01)
isolated from everyone * migr.b.	-0.29			
	(0.39)			
isolated from majority students * migr.b.		-0.17		
		(0.19)		
avoidance from majority * migr.b.			-0.01	
			(0.32)	
avoidance from minority * migr.b.				-0.13
				(0.22)
intercept	-1.16	-1.21	-1.15	-1.14
	(0.76)	(0.77)	(0.76)	(0.76)
adj. R^2	0.31	0.31	0.31	0.31
N	8376	8376	8376	8376

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Abbreviations: migr.b. = migration background.

Table A.8: Linear regression of grade point average (standardised), with migration background and social exclusion interaction effects, Germany.

	Model 1	Model 2	Model 3	Model 4
migration background (dummy)	-0.18**	-0.18**	-0.24***	-0.19**
	(0.07)	(0.07)	(0.07)	(0.07)
gender (female)	0.07*	0.07*	0.07*	0.07*
	(0.04)	(0.04)	(0.04)	(0.04)
age (in months)	-0.02***	-0.02***	-0.02***	-0.02***
	(0.00)	(0.00)	(0.00)	(0.00)
cognitive ability	0.05***	0.05***	0.05***	0.05***
	(0.00)	(0.00)	(0.00)	(0.00)
highest parental degree (ref: primary school)				
lower secondary school	-0.09	-0.09	-0.09	-0.09
	(0.08)	(0.08)	(0.08)	(0.08)
upper secondary school	-0.02	-0.02	-0.03	-0.02
	(0.08)	(0.08)	(0.08)	(0.08)
university degree	0.22*	0.22*	0.22*	0.22*
	(0.09)	(0.09)	(0.09)	(0.09)
parental unemployment	0.03	0.03	0.03	0.03
	(0.04)	(0.04)	(0.04)	(0.04)
other language spoken at home	0.11	0.11	0.11	0.11
	(0.06)	(0.06)	(0.06)	(0.06)
school resistance	-0.06***	-0.06***	-0.06***	-0.06***
	(0.01)	(0.01)	(0.01)	(0.01)
isolated from everyone	-0.06	-0.10	-0.09	-0.10
	(0.42)	(0.35)	(0.35)	(0.35)
isolated from majority students	-0.24***	-0.19	-0.25***	-0.24***
	(0.06)	(0.23)	(0.06)	(0.06)
avoidance by majority	-0.27**	-0.27**	-0.50***	-0.27**
	(0.10)	(0.10)	(0.14)	(0.10)
avoidance by minority	-0.06	-0.06	-0.05	-0.07
	(0.08)	(0.08)	(0.08)	(0.08)
number of friends (indegree)	0.02	0.02	0.01	0.02
	(0.01)	(0.01)	(0.01)	(0.01) v
isolated from everyone * migr.b.	-0.13			
	(0.77)			
isolated from majority students * migr.b.		-0.06		
		(0.24)		
avoidance from majority * migr.b.			0.43*	
			(0.18)	
avoidance from minority * migr.b.				0.04
				(0.19)
intercept	2.80***	2.81***	2.82***	2.81***
	(0.43)	(0.43)	(0.43)	(0.43)
adj. R^2	0.15	0.15	0.15	0.15
N	8376	8376	8376	8376

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

All models control for school track.

Abbreviations: migr.b. = migration background.

Table A.9: Linear regression of grade point average (standardised), with migration background and social exclusion interaction effects, Sweden.

	Model 1	Model 2	Model 3	Model 4
migration background	-0.03	-0.02	-0.04	-0.05
	(0.06)	(0.06)	(0.07)	(0.07)
gender (female)	0.14***	0.14***	0.14***	0.14***
	(0.03)	(0.03)	(0.03)	(0.03)
age (in months)	-0.01	-0.01	-0.01	-0.01
	(0.00)	(0.00)	(0.00)	(0.00)
cognitive ability	0.07***	0.07***	0.07***	0.07***
	(0.00)	(0.00)	(0.00)	(0.00)
highest parental degree (ref: primary school)				
lower secondary school	0.04	0.04	0.04	0.05
	(0.13)	(0.13)	(0.13)	(0.13)
upper secondary school	0.20*	0.20*	0.19*	0.20*
	(0.08)	(0.08)	(0.08)	(0.08)
university degree	0.53***	0.53***	0.53***	0.53***
	(0.08)	(0.08)	(0.08)	(0.08)
parental unemployment	-0.05	-0.06	-0.05	-0.05
	(0.05)	(0.05)	(0.05)	(0.05)
other language spoken at home	0.03	0.03	0.03	0.03
	(0.06)	(0.06)	(0.06)	(0.06)
school resistance	-0.07***	-0.07***	-0.07***	-0.07***
	(0.01)	(0.01)	(0.01)	(0.01)
isolated from everyone	-0.11	-0.08	-0.09	-0.08
	(0.36)	(0.27)	(0.27)	(0.27)
isolated from majority students	-0.01	0.06	-0.02	-0.01
	(0.06)	(0.19)	(0.06)	(0.06)
avoidance by majority	-0.53***	-0.52***	-0.61**	-0.53***
	(0.14)	(0.14)	(0.20)	(0.14)
avoidance by minority	-0.12	-0.11	-0.12	-0.13
	(0.06)	(0.06)	(0.06)	(0.07)
number of friends (indegree)	0.02*	0.02*	0.02*	0.02*
	(0.01)	(0.01)	(0.01)	(0.01)
isolated from everyone * migr.b.	0.06			
	(0.54)			
isolated from majority students * migr.b.		-0.09		
		(0.20)		
avoidance from majority * migr.b.			0.14	
			(0.26)	
avoidance from minority * migr.b.				0.14
				(0.24)
intercept	0.08	0.08	0.09	0.08
	(0.68)	(0.68)	(0.68)	(0.68)
adj. R^2	0.26	0.26	0.26	0.26
N	8376	8376	8376	8376

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Abbreviations: migr.b. = migration background.

Table A.10: Linear regression of grade point average (standardised), with country dummy interaction effects.

	Model 1	Model 2	Model 3	Model 4
migration background (ref: native)				
newly arrived	-0.21**	-0.22**	-0.21**	-0.21**
	(0.07)	(0.07)	(0.07)	(0.07)
1st generation	-0.12**	-0.13**	-0.12**	-0.12**
	(0.05)	(0.05)	(0.05)	(0.05)
2nd generation	-0.06	-0.06	-0.06	-0.06
	(0.04)	(0.04)	(0.04)	(0.04)
isolated from everyone	-0.14	-0.12	-0.11	-0.13
	(0.21)	(0.16)	(0.16)	(0.16)
isolated from natives	-0.09**	-0.05	-0.09*	-0.09**
	(0.04)	(0.06)	(0.04)	(0.04)
avoidance by majority	-0.32***	-0.32***	-0.27	-0.32***
	(0.07)	(0.07)	(0.15)	(0.07)
avoidance by minority	-0.03	-0.03	-0.04	0.02
	(0.04)	(0.04)	(0.04)	(0.06)
Survey country (ref: England)				
Germany	-0.06*	-0.04	-0.07*	-0.05
	(0.03)	(0.03)	(0.03)	(0.03)
Sweden	-0.13***	-0.13***	-0.11***	-0.11***
	(0.03)	(0.03)	(0.03)	(0.03)
number of friends (indegree)	0.02***	0.02***	0.02***	0.02***
	(0.01)	(0.01)	(0.01)	(0.01)
isolated from everyone * Germany	0.14			
	(0.39)			
isolated from everyone * Sweden	-0.02			
	(0.32)			
isolated from natives * Germany		-0.16*		
		(0.08)		
isolated from natives * Sweden		0.03		
		(0.08)		
avoidance by majority * Germany			0.06	
			(0.18)	
avoidance by majority * Sweden			-0.27	
			(0.20)	
avoidance by minority * Germany				-0.02
				(0.10)
avoidance by minority * Sweden				-0.12
				(0.09)
intercept	0.97**	0.90**	0.98**	0.95**
	(0.32)	(0.32)	(0.32)	(0.32)
adj. R^2	0.21	0.22	0.22	0.21
N	8376	8376	8376	8376

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

All models control for gender, age, cognitive ability, highest parental degree, parental unemployment, other language spoken at home, and school resistance.

Appendix B

Appendix Chapter 3

Table B.1: Logistic regression predicting being in the analytical sample, reporting average marginal effects.

	Model
the Netherlands (ref. Germany)	0.08**
	(0.03)
Sweden (ref. Germany)	0.13***
	(0.03)
Individual univ. aspirations wave 1	0.02
	(0.01)
Individual univ. aspirations wave 2	0.05***
	(0.01)
Sex (female)	0.02
	(0.01)
Age (in months)	-0.00***
	(0.00)
Migration background (binary)	-0.03*
	(0.01)
Cognitive ability index	0.01***
	(0.00)
At least one parent has a univ. degree	0.02
	(0.01)
Sample stratum	-0.02
	(0.01)
Num. obs.	11172
Log Likelihood	-6993.04

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table B.2: Logistic regression predicting individual university aspirations at wave 2 (reporting average marginal effects, standard errors clustered on class level). Friends as the social milieu. Non-imputed data.

	Model 1	Model 2	Model 3	Model 4
friends: av. univ. aspirations	0.49***	0.18***	0.12***	0.12***
	(0.02)	(0.02)	(0.02)	(0.03)
individual univ. aspirations, w1		0.39***	0.38***	0.38***
		(0.01)	(0.01)	(0.01)
gender (female)		0.06***	0.06***	0.06***
		(0.01)	(0.01)	(0.01)
age (in months)		-0.01***	-0.00***	-0.00***
		(0.00)	(0.00)	(0.00)
migration background (vs. native)		0.09***	0.09***	0.09***
		(0.02)	(0.02)	(0.03)
cognitive ability index		0.01***	0.01***	0.01***
		(0.00)	(0.00)	(0.00)
at least one parent has a univ. degree		0.07***	0.06***	0.06***
		(0.02)	(0.02)	(0.02)
parental univ. aspirations		0.16***	0.15***	0.15***
		(0.02)	(0.02)	(0.02)
social milieu: friends				
av. cognitive ability		-0.05	-0.05	-0.05
		(0.03)	(0.03)	(0.03)
prop. natives		0.01***	0.01**	0.01**
		(0.00)	(0.00)	(0.00)
prop. parental degree		0.14***	0.13***	0.13***
		(0.03)	(0.03)	(0.03)
school class				
av. univ. aspirations			0.35***	0.35***
			(0.06)	(0.06)
prop. natives			0.09	0.09
			(0.06)	(0.06)
av. cognitive ability			-0.00	-0.00
			(0.00)	(0.00)
prop. parental degree			-0.06	-0.06
			(0.07)	(0.07)
friends: av. univ. aspirations * migr.b.				-0.01
				(0.05)
Num. obs.	7147	7147	7147	7147
Log Likelihood	-4071.69	-3191.04	-3164.98	-3164.97

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

All models control for survey country and sampling stratum.

Abbreviations: migr.b. = migration background.

Table B.3: Logistic regression predicting individual university aspirations at wave 2 (reporting average marginal effects, standard errors clustered on class level). Friendship group as the social milieu. Non-imputed data.

	Model 1	Model 2	Model 3	Model 4
fr. group: av. univ. aspirations	0.39***	0.17***	0.12***	0.11***
	(0.02)	(0.02)	(0.02)	(0.02)
individual univ. aspirations, w1		0.39***	0.38***	0.38***
		(0.01)	(0.01)	(0.01)
gender (female)		0.06***	0.06***	0.06***
		(0.01)	(0.01)	(0.01)
age (in months)		-0.01***	-0.00***	-0.00***
		(0.00)	(0.00)	(0.00)
migration background (vs. native)		0.10***	0.09***	0.09***
		(0.02)	(0.02)	(0.02)
cognitive ability index		0.01***	0.01***	0.01***
		(0.00)	(0.00)	(0.00)
at least one parent has a univ. degree		0.07***	0.06***	0.06***
		(0.02)	(0.02)	(0.02)
parental univ. aspirations		0.16***	0.15***	0.15***
		(0.02)	(0.02)	(0.02)
social milieu: friendship group				
av. cognitive ability		0.01***	0.01***	0.01***
		(0.00)	(0.00)	(0.00)
prop. parental degree		0.08**	0.06*	0.06*
		(0.02)	(0.02)	(0.02)
prop. natives			-0.03	-0.03
			(0.02)	(0.02)
school class				
av. univ. aspirations			0.26***	0.26***
			(0.05)	(0.05)
prop. natives			0.05	0.05
			(0.04)	(0.04)
av. cognitive ability			-0.00	-0.00
			(0.00)	(0.00)
prop. parental degree			0.01	0.01
			(0.06)	(0.06)
fr.group: av. univ. aspirations * migr.b.				0.01
				(0.04)
Num. obs.	7147	7147	7147	7147
Log Likelihood	-4128.66	-3196.87	-3169.64	-3169.62

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

All models control for survey country and sampling stratum.

Abbreviations: fr.group = friendship group, migr.b. = migration background

Table B.4: Logistic regression predicting individual univ. aspirations at wave 2 (reporting average marginal effects, standard errors clustered on class level). Friends as the social milieu. *German sub-sample.*

	Model 1	Model 2	Model 3	Model 4
friends: av. univ. aspirations	0.23***	0.09	0.08	0.07
	(0.05)	(0.06)	(0.06)	(0.07)
individual univ. aspirations, w1		0.43***	0.43***	0.43***
		(0.03)	(0.03)	(0.03)
gender (female)		0.00	0.01	0.01
		(0.03)	(0.03)	(0.03)
age (in months)		-0.01**	-0.01**	-0.01**
		(0.00)	(0.00)	(0.00)
migration background (vs. native)		0.08**	0.09**	0.08
		(0.03)	(0.03)	(0.05)
cognitive ability index		0.00	0.00	0.00
		(0.00)	(0.00)	(0.00)
at least one parent has a univ. degree		0.15***	0.16***	0.16***
		(0.04)	(0.04)	(0.04)
parental univ. aspirations		0.25***	0.24***	0.24***
		(0.04)	(0.04)	(0.04)
social milieu: friends				
av. cognitive ability		-0.03	-0.04	-0.03
		(0.05)	(0.06)	(0.06)
prop. natives		-0.00	-0.00	-0.00
		(0.01)	(0.01)	(0.01)
prop. parental degree		0.11	0.15	0.15
		(0.07)	(0.08)	(0.08)
school class				
av. univ. aspirations			0.28	0.29
			(0.16)	(0.16)
prop. natives			-0.00	-0.01
			(0.11)	(0.11)
av. cognitive ability			-0.02*	-0.02*
			(0.01)	(0.01)
prop. parental degree			-0.23	-0.23
			(0.16)	(0.16)
friends: av. univ. aspirations * migr.b.				0.03
				(0.09)
Num. obs.	2235	2235	2235	2235
Log Likelihood	-1300.66	-1030.54	-1026.08	-1026.04

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

All models control for sampling stratum and school type (track).

Abbreviations: migr.b. = migration background, univ. = university

Table B.5: Logistic regression predicting individual univ. aspirations at wave 2 (reporting average marginal effects, standard errors clustered on class level). Friends as the social milieu. *Dutch sub-sample.*

	Model 1	Model 2	Model 3	Model 4
friends: av. univ. aspirations	0.23***	0.06	0.04	0.04
	(0.04)	(0.04)	(0.04)	(0.05)
individual univ. aspirations, w1		0.35***	0.35***	0.35***
		(0.02)	(0.02)	(0.02)
gender (female)		0.01	0.01	0.01
		(0.02)	(0.02)	(0.02)
age (in months)		-0.00	-0.00	-0.00
		(0.00)	(0.00)	(0.00)
migration background (vs. native)		0.14***	0.14***	0.13**
		(0.03)	(0.03)	(0.05)
cognitive ability index		-0.00	-0.00	-0.00
		(0.00)	(0.00)	(0.00)
at least one parent has a univ. degree		-0.03	-0.03	-0.03
		(0.03)	(0.03)	(0.03)
parental univ. aspirations		0.11***	0.11***	0.11***
		(0.03)	(0.03)	(0.03)
social milieu: friends				
av. cognitive ability		-0.07	-0.07	-0.07
		(0.04)	(0.05)	(0.05)
prop. natives		0.02***	0.01**	0.01**
		(0.01)	(0.01)	(0.01)
prop. parental degree		0.16**	0.16**	0.16**
		(0.06)	(0.06)	(0.06)
school class				
av. univ. aspirations			0.09	0.09
			(0.12)	(0.12)
prop. natives			0.03	0.03
			(0.10)	(0.10)
av. cognitive ability			0.01*	0.01*
			(0.01)	(0.01)
prop. parental degree			0.03	0.03
			(0.14)	(0.15)
friends: av. univ. aspirations * migr.b.				0.02
				(0.10)
Num. obs.	2101	2101	2101	2101
Log Likelihood	-1135.49	-935.96	-932.20	-932.17

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

All models control for sampling stratum and school type (track).

Abbreviations: migr.b. = migration background, univ. = university

Table B.6: Logistic regression predicting individual univ. aspirations at wave 2 (reporting average marginal effects, standard errors clustered on class level). Friends as the social milieu. *Swedish sub-sample.*

	Model 1	Model 2	Model 3	Model 4
friends: av. univ. aspirations	0.27***	0.09***	0.09***	0.10**
	(0.03)	(0.03)	(0.03)	(0.03)
individual univ. aspirations, w1		0.29***	0.29***	0.29***
		(0.02)	(0.02)	(0.02)
gender (female)		0.09***	0.09***	0.09***
		(0.02)	(0.02)	(0.02)
age (in months)		-0.00	-0.00	-0.00
		(0.00)	(0.00)	(0.00)
migration background (vs. native)		0.05**	0.05**	0.08*
		(0.02)	(0.02)	(0.03)
cognitive ability index		0.01***	0.01***	0.01***
		(0.00)	(0.00)	(0.00)
at least one parent has a univ. degree		0.05***	0.05***	0.05***
		(0.01)	(0.01)	(0.01)
parental univ. aspirations		0.09***	0.09***	0.09***
		(0.02)	(0.02)	(0.02)
social milieu: friends				
av. cognitive ability		-0.03	-0.04	-0.04
		(0.03)	(0.03)	(0.03)
prop. natives		0.00	0.00	0.00
		(0.00)	(0.00)	(0.00)
prop. parental degree		0.07*	0.06*	0.06*
		(0.03)	(0.03)	(0.03)
school class				
av. univ. aspirations			0.20**	0.20**
			(0.08)	(0.08)
prop. natives			0.13	0.13
			(0.07)	(0.07)
av. cognitive ability			-0.00	-0.00
			(0.00)	(0.00)
prop. parental degree			-0.05	-0.05
			(0.07)	(0.07)
friends: av. univ. aspirations * migr.b.				-0.05
				(0.05)
Num. obs.	2811	2811	2811	2811
Log Likelihood	-1465.69	-1159.44	-1154.35	-1153.96

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

All models control for sampling stratum.

Abbreviations: migr.b. = migration background, univ. = university

Table B.7: Logistic regression predicting individual univ. aspirations at wave 2 (reporting average marginal effects, standard errors clustered on class level). Friendship group as the social milieu. *German sub-sample.*

	Model 1	Model 2	Model 3	Model 4
fr. group: av. univ. aspirations	0.15***	0.09	0.07	0.06
	(0.04)	(0.05)	(0.05)	(0.06)
individual univ. aspirations, w1		0.43***	0.43***	0.43***
		(0.03)	(0.03)	(0.03)
gender (female)		0.00	0.00	0.00
		(0.03)	(0.03)	(0.03)
age (in months)		-0.01**	-0.01**	-0.01**
		(0.00)	(0.00)	(0.00)
migration background (vs. native)		0.08**	0.08**	0.07
		(0.03)	(0.03)	(0.04)
cognitive ability index		0.00	0.00	0.00
		(0.00)	(0.00)	(0.00)
at least one parent has a univ. degree		0.15***	0.16***	0.16***
		(0.04)	(0.04)	(0.04)
parental univ. aspirations		0.24***	0.24***	0.24***
		(0.04)	(0.04)	(0.04)
social milieu: friends				
fr.group: av. cognitive ability		-0.00	0.00	0.00
		(0.00)	(0.01)	(0.01)
fr.group: prop. parental degree		0.08	0.11	0.11
		(0.05)	(0.06)	(0.06)
fr.group: prop. natives			-0.02	-0.02
			(0.04)	(0.04)
school class				
class: av. univ. aspirations			0.24*	0.24*
			(0.11)	(0.11)
class: prop. natives			-0.00	-0.00
			(0.09)	(0.09)
class: av. cognitive ability			-0.01	-0.01
			(0.01)	(0.01)
class: prop. parental degree			-0.15	-0.15
			(0.12)	(0.12)
fr. group: av. univ. aspirations * migr.b.				0.03
				(0.08)
Num. obs.	2235	2235	2235	2235
Log Likelihood	-1304.95	-1030.99	-1027.48	-1027.41

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

All models control for sampling stratum and school type (track). Abbreviations: fr.group =friendship group, migr.b. = migration background, univ. = university

Table B.8: Logistic regression predicting individual univ. aspirations at wave 2 (reporting average marginal effects, standard errors clustered on class level). Friendship group as the social milieu. *Dutch sub-sample.*

	Model 1	Model 2	Model 3	Model 4
fr. group: av. univ. aspirations	0.17***	0.08*	0.06	0.05
	(0.03)	(0.03)	(0.03)	(0.04)
individual univ. aspirations, w1		0.35***	0.35***	0.35***
		(0.02)	(0.02)	(0.02)
gender (female)		0.02	0.02	0.02
		(0.02)	(0.02)	(0.02)
age (in months)		-0.00	-0.00	-0.00
		(0.00)	(0.00)	(0.00)
migration background (vs. native)		0.15***	0.14***	0.13***
		(0.02)	(0.03)	(0.04)
cognitive ability index		-0.00	-0.00	-0.00
		(0.00)	(0.00)	(0.00)
at least one parent has a univ. degree		-0.02	-0.03	-0.03
		(0.03)	(0.03)	(0.03)
parental univ. aspirations		0.12***	0.11***	0.11***
		(0.03)	(0.03)	(0.03)
social milieu: friends				
fr.group: av. cognitive ability		0.01**	0.01*	0.01*
		(0.00)	(0.00)	(0.00)
fr.group: prop. parental degree		0.10*	0.08	0.09
		(0.04)	(0.04)	(0.04)
fr.group: prop. natives			-0.02	-0.02
			(0.04)	(0.04)
school class				
class: av. univ. aspirations			-0.01	-0.01
			(0.09)	(0.09)
class: prop. natives			-0.06	-0.05
			(0.08)	(0.08)
class: av. cognitive ability			0.02**	0.02**
			(0.01)	(0.01)
class: prop. parental degree			0.10	0.10
			(0.12)	(0.12)
fr. group: av. univ. aspirations * migr.b.				0.02
				(0.08)
Num. obs.	2101	2101	2101	2101
Log Likelihood	-1139.55	-940.76	-935.47	-935.44

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

All models control for sampling stratum and school type (track). Abbreviations: fr.group = friendship group, migr.b. = migration background, univ. = university

Table B.9: Logistic regression predicting individual univ. aspirations at wave 2 (reporting average marginal effects, standard errors clustered on class level). Friendship group as the social milieu. *Swedish sub-sample.*

	Model 1	Model 2	Model 3	Model 4
fr. group: av. univ. aspirations	0.22***	0.10***	0.09***	0.09***
	(0.03)	(0.02)	(0.02)	(0.03)
individual univ. aspirations, w1		0.29***	0.29***	0.29***
		(0.02)	(0.02)	(0.02)
gender (female)		0.09***	0.09***	0.09***
		(0.02)	(0.02)	(0.02)
age (in months)		-0.00	-0.00	-0.00
		(0.00)	(0.00)	(0.00)
migration background (vs. native)		0.05**	0.05**	0.06*
		(0.02)	(0.02)	(0.03)
cognitive ability index		0.01***	0.01***	0.01***
		(0.00)	(0.00)	(0.00)
at least one parent has a univ. degree		0.05***	0.05***	0.05***
		(0.01)	(0.01)	(0.01)
parental univ. aspirations		0.10***	0.09***	0.09***
		(0.02)	(0.02)	(0.02)
social milieu: friends				
fr.group: av. cognitive ability		0.01*	0.01*	0.01*
		(0.00)	(0.00)	(0.00)
fr.group: prop. parental degree		0.02	0.02	0.02
		(0.02)	(0.02)	(0.02)
fr.group: prop. natives			-0.04	-0.04
			(0.03)	(0.03)
school class				
class: av. univ. aspirations			0.14*	0.14*
			(0.06)	(0.06)
class: prop. natives			0.10	0.10
			(0.05)	(0.05)
class: av. cognitive ability			-0.01	-0.01
			(0.00)	(0.00)
class: prop. parental degree			0.01	0.01
			(0.05)	(0.05)
fr. group: av. univ. aspirations * migr.b.				-0.02
				(0.05)
Num. obs.	2811	2811	2811	2811
Log Likelihood	-1472.25	-1157.56	-1150.85	-1150.79

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

All models control for sampling stratum and school type (track). Abbreviations: fr.group = friendship group, migr.b. = migration background, univ. = university

Appendix C

Appendix Chapter 4

Table C.1: Variable descriptives, analytical and full sample.

	analytical sample						*full sample*					
	mean	SD	median	min	max	N	mean	SD	median	min	max	N
Cognitive ability test	17.54	4.94	18	0	27	4189	17.51	4.97	18	0	27	4773
Gender: female						2518						2744
Gender: male						2460						2683
Students having a STEM-favourite subject, wave 1						862						959
Students having a STEM-favourite subject, wave 2						687						732
Parental education: no school leaving certificate						173						192
degree from lower secondary school						164						179
degree from upper secondary school						1480						1611
university degree						2055						2241

Table C.2: Network descriptives, included and dropped groups.

	Groups included in analysis				*Dropped groups*			
	min	max	mean	sd	min	max	mean	sd
friendship networks								
size	9.00	34.00	22.86	3.67	6.00	27.00	13.55	4.60
density	0.00	0.24	0.13	0.04	0.00	0.47	0.16	0.10
degrees	0.00	5.00	2.74	1.85	0.00	5.00	1.89	1.76
avgDegree	0.00	4.70	2.71	0.73	0.00	3.90	1.81	1.03
clustering	0.25	0.88	0.57	0.11	0.00	1.00	0.61	0.24
jaccard	0.00	0.58	0.33	0.11	0.00	0.83	0.27	0.22
same gender ties (t1)	0.54	1.00	0.90	0.08	0.65	1.00	0.91	0.08
same gender ties (t2)	0.51	1.00	0.87	0.10	0.63	1.00	0.89	0.11
subject networks								
density	0.03	0.08	0.05	0.01	0.00	0.08	0.05	0.02
degrees	0.00	2.00	0.89	0.42	0.00	2.00	0.82	0.52
avgDegree	0.48	1.39	0.89	0.13	0.00	1.33	0.83	0.27
jaccard	0.05	0.56	0.27	0.09	0.00	0.45	0.15	0.13

Estimation procedure

For the estimation, we used a Bayesian MCMC procedure (Gelman et al., 2013). Throughout the process, we followed the guidelines of Snijders and Koskinen (2017). The model includes 460 parameters for the 218 groups: 436 rate parameters, and 24 global parameters. Rate parameters, which capture how much change an average actor can make in his or her network, are different for each group and network: therefore, with the 218 groups of one-mode and two-mode networks, we had 218 rate parameters for the friendship network, and 218 for the favourite subject network. Of the global parameters, we specified fixed and random effects. A fixed effect is the same for all groups; for a random effect, the parameter varies between groups following a multivariate normal distribution. To obtain a good statistical power, and because we did not specifically assume differences between our groups, all parameters playing a role in our hypothesis tests were modelled as fixed. In our main model, 13 global effects were fixed, and 11 were random. The full table of results in Table A.4 in the Appendix also provides the list of all global parameters and whether they were random (+) or fixed (-).

Since we applied a Bayesian analysis, we had to specify prior distributions for our parameters. The priors for rate parameters were data-dependent, which is the default in the estimation procedure. While priors are better specified in the non-default way for a small number of group, this does not apply to our case, since we have a comparably very large number of groups. The default priors that were used for all parameters, random and fixed, were 0, with a standard deviation of 1.

Table C.3: Details on model specification (friendship network part).

Dependent variable: friendship (ego creates or maintains a tie to alter)

Effect name (*RSiena effect name*)	The tendency is modelled...
Outdegree (*density*)	... to create and maintain friendships.
Reciprocity (*recip*)	... to reciprocate friendships.
Transitive triplets (*transTrip*)	... to create and maintain friendships to friends of friends.
Transitive reciprocated triplets (*transRecTrip*)	... to reciprocate ties to friends of friends.
Alter's popularity, sqrt (*inPopSqrt*)	... to create and maintain ties to actors that are popular in the network.
Ego's popularity (sqrt) (*inActSqrt*)	... to create and maintain ties to actors that have ties to more others.
Ego's activity, sqrt (*outAct*)	... for actors who already have more ties, to create and maintain more
Female alter (*altX*)	... to create and maintain more ties to girls than to boys
Female ego (*egoX*)	... for girls to create and maintain more ties
Same gender of ego and alter (*sameX*)	... for actors to be tied to those of the Same gender of ego and alter
Shared favourite subject of ego and alter (*from*)	... for actors that have the same favourite subject to become friends (friend selection)
Female ego * shared favourite subject of ego and alter (*egoX * from*)	... for girls to behave differently regarding friend selection based on favourite subjects

Notes. P = parameter. The covariate for gender is non-centred. Cognitive ability is ranging from 0 to 27 and is centred. Parental education ranges from 1 to 4, and is non-centred. STEM is a binary covariate, with 1 for STEM and 0 for non-STEM.

Table C.4: Details on model specification (favourite subject network part).

Dependent variable: favourite subject (ego creates or maintains a tie to a subject)

Effect name (*RSiena effect name*)	The tendency is modelled...
Outdegree (*density*)	... to name favourite subjects
Subject popularity, sqrt (*inPopSqrt*)	... for more popular subjects to attract more ties
Ego's popularity, sqrt (*inPopIntn*, p=2)	... for actors who are nominated by more others as friends to name more favourite subjects
Ego's friendship activity, sqrt (*outActIntn*, p=2)	... for actors who nominate more others in the friendship network to name more favourite subjects
Female ego (*egoX*)	... for girls to name more favourite subjects
Friend influence (*to*)	... to start or keep liking a subject if more friends likes it (friend influence)
Female ego * friend influence (*to * egoX*)	... for girls to behave differently regarding friend influence on favourite subjects
STEM subject (*altX*)	... to name STEM subjects
Female ego * STEM subject (*egoX * altX*)	... for girls to behave differently from boys in naming STEM subjects
Subject popularity among females (*inPopX*, p=2)	... for subjects that are more popular among girls to attract more ties
Subject popularity among males (*inPopX*, p=2)	... for subjects that are more popular among boys to attract more ties
Female ego * subject popularity among females (*egoX * inPopX*)	... for subjects that are more popular among girls to attract more ties from girls
Female ego * subject popularity among males (*egoX * inPopX*)	... for subjects that are more popular among boys to attract more ties from girls
STEM subject * subject popularity among females (*altX * inPopX*)	... for subjects that are more popular among girls to attract more ties if they are STEM subjects
STEM subject * subject popularity among males (*altX * inPopX*)	... for subjects that are more popular among boys to attract more ties if they are STEM subjects
Female ego * STEM subject * subject popularity among females (*egoX * altX * inPopX*)	... for subjects that are more popular among girls to attract more ties from girls if they are STEM subjects
Female ego * STEM subject * subject popularity among males (*egoX * altX * inPopX*)	... for subjects that are more popular among boys to attract more ties from girls if they are STEM subjects
Parental education ego (*egoX*)	... for actors with higher parental education to name more favourite subjects
Cognitive ability ego (*egoX*)	... for actors with higher cognitive ability to name more favourite subjects
parental education ego * STEM subject (*egoX * altX*)	... for actors with higher parental education to name STEM subjects
cognitive ability ego * STEM subject (*egoX * altX*)	... for actors with higher cognitive ability to name STEM subjects

Notes. P = parameter. The covariate for gender is non-centred. Cognitive ability is ranging from 0 to 27 and is centred. Parental education ranges from 1 to 4, and is non-centred. STEM is a binary covariate, with 1 for STEM and 0 for non-STEM.

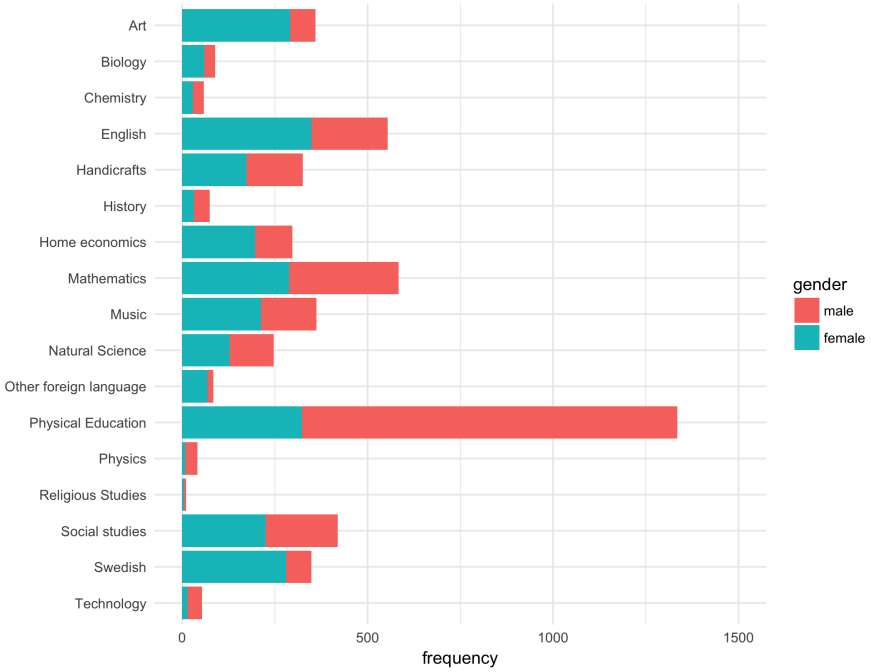

Figure C.1: Subject popularity by gender, wave 1.

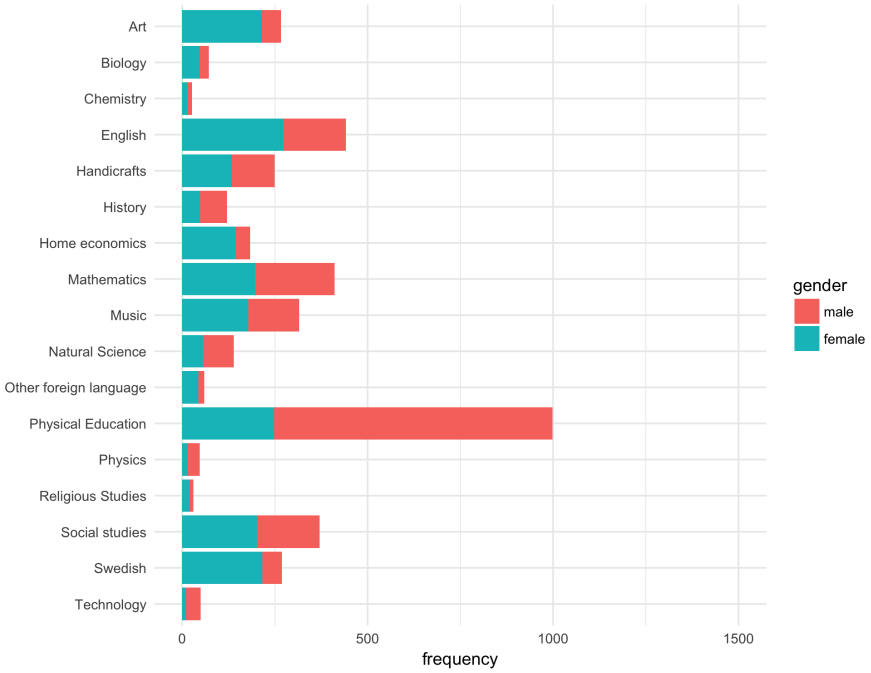

Figure C.2: Subject popularity by gender, wave 2.

Table C.5: Proportion of students having a STEM-favourite subjects by gender, for all four countries that are included in *CILS4EU*.

		wave 1	wave 2
England	male	33.70%	40.00%
	female	21.40%	26.00%
Germany	male	31.40%	33.90%
	female	19.60%	20.50%
Netherlands	male	18.80%	21.10%
	female	13.70%	12.00%
Sweden	male	20.60%	19.60%
	female	18.60%	14.50%

Table C.6: Results, random-coefficient multilevel SAOMs. Main model, both peer effects, with covariate controls.

	β	SD	varying
Friendship part			
Outdegree (density)	-2.292***	(0.134)	+
Alter's popularity	0.002	(0.048)	+
Ego's popularity	0.139	(0.086)	+
Ego's activity	-0.495***	(0.061)	+
Reciprocity	2.243***	(0.070)	+
Transitive triplets	0.794***	(0.034)	+
Transitive reciprocated triplets	-0.482***	(0.041)	+
Female ego	0.040	(0.033)	-
Female alter	-0.019	(0.028)	-
Same gender of ego and alter	0.493***	(0.027)	-
Shared favourite subject of ego and alter	-0.303***	(0.102)	-
Female ego * shared favourite subject of ego and alter	0.259***	(0.093)	-
Favourite subject part			
Outdegree (density)	-3.083***	(0.245)	+
Subject popularity	0.120	(0.101)	+
Ego's popularity	-0.056	(0.070)	+
Ego's friendship activity	-0.277***	(0.075)	+
Female ego	0.643***	(0.147)	-
Subject popularity among females	0.331*	(0.141)	-
STEM subject	0.324	(0.272)	-
Female ego * subject popularity among females	-0.161	(0.100)	-
Female ego * STEM subject	-0.946***	(0.190)	-
STEM subject * subject popularity among females	-0.018	(0.163)	-
Female ego * STEM subject * subject popularity among females	0.353***	(0.139)	-
Friend influence	1.089***	(0.091)	-
Female ego * friend influence	-0.477***	(0.092)	-
Parental education ego	0.013	(0.024)	-
Parental education ego * STEM subject	-0.020	(0.032)	-
Cognitive ability ego	-0.027***	(0.008)	-
Cognitive ability ego * STEM subject	0.053***	(0.012)	-

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table C.7: Results, random-coefficient multilevel SAOMs. Main model, both peer effects, with opposite sex peer exposure, with covariate controls.

	β	SD	varying
Friendship part			
Outdegree (density)	-2.28***	(0.126)	+
Alter's popularity	0.001	(0.045)	+
Ego's popularity	0.112	(0.086)	+
Ego's activity	-0.481***	(0.070)	+
Reciprocity	2.240***	(0.069)	+
Transitive triplets	0.790***	(0.034)	+
Transitive reciprocated triplets	-0.478***	(0.041)	+
Female ego	0.050	(0.038)	-
Female alter	-0.016	(0.025)	-
Same gender of ego and alter	0.487***	(0.023)	-
Shared favourite subject of ego and alter	-0.207***	(0.093)	-
Female ego * shared favourite subject of ego and alter	0.207***	(0.093)	-
Favourite subject part			
Outdegree (density)	-2.75***	(0.207)	+
Ego's popularity	-0.064	(0.062)	+
Ego's friendship activity	-0.231***	(0.075)	+
Female ego	0.572***	(0.138)	-
STEM subject	0.089	(0.246)	-
Subject popularity among females	-0.042	(0.131)	-
Subject popularity among males	0.517***	(0.118)	-
Female ego * STEM subject	-0.876***	(0.193)	-
Female ego * subject popularity among females	0.130	(0.114)	-
Female ego * subject popularity among males	-0.282	(0.100)	-
STEM subject * subject popularity among females	0.301	(0.215)	-
STEM subject * subject popularity among males	-0.253	(0.210)	-
Female ego * STEM subject * subject popularity among females	0.158***	(0.188)	-
Female ego * STEM subject * subject popularity among males	0.130***	(0.188)	-
Friend influence	0.820***	(0.088)	-
Female ego * friend influence	-0.310***	(0.089)	-
Parental education ego	-0.021	(0.021)	-
Parental education ego * STEM subject	0.007	(0.037)	-
Cognitive ability ego	-0.026***	(0.008)	-
Cognitive ability ego * STEM subject	0.050***	(0.012)	-

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table C.8: Results, random-coefficient multilevel SAOMs. Main model, no gender.

	β	SD	varying
Friendship part			
Outdegree (density)	-1.756***	(0.118)	+
Alter's popularity	-0.063	(0.046)	+
Ego's popularity	0.105	(0.077)	+
Ego's activity	-0.535***	(0.058)	+
Reciprocity	2.330***	(0.064)	+
Transitive triplets	0.857***	(0.034)	+
Transitive reciprocated triplets	-0.502***	(0.039)	+
Shared favourite subject of ego and alter	-0.017	(0.045)	-
Favourite subject part			
Outdegree (density)	-3.046***	(0.086)	+
Subject popularity	0.576***	(0.054)	+
Ego's popularity	-0.072	(0.061)	+
Ego's friendship activity	-0.278***	(0.071)	+
Friend influence	0.633***	(0.044)	-

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table C.9: Results, random-coefficient multilevel SAOMs. Main model, no gender, with subject effects.

	β	SD	varying
Friendship part			
Outdegree (density)	-1.764***	(0.123)	+
Alter's popularity	-0.060	(0.046)	+
Ego's popularity	0.119	(0.089)	+
Ego's activity	-0.546***	(0.063)	+
Reciprocity	2.344***	(0.065)	+
Transitive triplets	0.860***	(0.034)	+
Transitive reciprocated triplets	-0.507***	(0.039)	+
Shared favourite subject of ego and alter	-0.032	(0.053)	-
Favourite subject part			
Outdegree (density)	-1.074***	(0.163)	+
Subject popularity	0.087	(0.062)	+
Ego's popularity	-0.055	(0.062)	+
Ego's friendship activity	-0.285***	(0.068)	+
Friend influence	0.649***	(0.041)	-
Arts alter	-0.310***	(0.090)	-
Biology alter	-1.187***	(0.158)	-
Chemistry alter	-2.015***	(0.203)	-
Crafts alter	-0.245***	(0.092)	-
History alter	-0.665***	(0.112)	-
Home economics alter	-0.590***	(0.098)	-
Maths alter	-0.144	(0.085)	-
Music alter	-0.239***	(0.087)	-
Natural Sciences alter	-0.796***	(0.111)	-
Foreign languages alter	-1.359***	(0.169)	-
Physical education alter	0.064	(0.080)	-
Physics alter	-1.431***	(0.150)	-
Religious studies alter	-1.906***	(0.192)	-
Social sciences alter	-0.051	(0.083)	-
Swedish alter	-0.294***	(0.090)	-
Technology alter	-1.449***	(0.155)	-

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table C.10: Results, random-coefficient multilevel SAOMs. Main model, both peer effects, with covariate controls and subject effects (*friendship part*).

	β	SD	varying
Friendship part			
Outdegree (density)	-2.292***	(0.139)	+
Alter's popularity	0.002	(0.049)	+
Ego's popularity	0.138	(0.091)	+
Ego's activity	-0.495***	(0.069)	+
Reciprocity	2.230***	(0.068)	+
Transitive triplets	0.789***	(0.037)	+
Transitive reciprocated triplets	-0.479***	(0.043)	+
Female ego	0.042	(0.025)	-
Female alter	-0.012	(0.028)	-
Same gender of ego and alter	0.495***	(0.029)	-
Shared favourite subject of ego and alter	-0.308***	(0.088)	-
Female ego * shared favourite subject of ego and alter	0.254*	(0.093)	-

***$p < 0.001$, **$p < 0.01$, *$p < 0.05$

Table C.11: Results, random-coefficient multilevel SAOMs. Main model, both peer effects, with covariate controls and subject effects (*favourite subject part*).

	β	SD	varying
Favourite subject part			
Outdegree (density)	-1.713***	(0.220)	+
Subject popularity	-0.097	(0.127)	+
Ego's popularity	-0.051	(0.069)	+
Ego's friendship activity	-0.273***	(0.078)	+
Female ego	0.775***	(0.123)	-
Subject popularity among females	0.332***	(0.124)	-
STEM subject			
Female ego * subject popularity among females	-0.267***	(0.084)	-
Female ego * STEM subject	-1.065***	(0.196)	-
STEM subject * subject popularity among females	-0.255	(0.153)	-
Female ego * STEM subject * subject popularity among females	0.445***	(0.118)	-
Friend influence	0.998***	(0.090)	-
Female ego * friend influence	-0.374***	(0.084)	-
Parental education ego	0.018	(0.019)	-
Parental education ego * STEM subject	-0.03	(0.033)	-
Cognitive ability ego	-0.028***	(0.007)	-
Cognitive ability ego * STEM subject	0.051***	(0.012)	-
Subject: Arts	-0.338***	(0.083)	-
Subject: Biology	-1.163***	(0.118)	-
Subject: Chemistry	-1.025***	(0.361)	-
Subject: Crafts	-0.294***	(0.077)	-
Subject: History	-0.719***	(0.130)	-
Subject: Home economics	-0.632***	(0.100)	-
Subject: Maths	0.785***	(0.301)	-
Subject: Music	-0.27***	(0.068)	-
Subject: Natural Sciences	0.171	(0.298)	-
Subject: Foreign languages	-1.422***	(0.155)	-
Subject: Physical education	0.045	(0.097)	-
Subject: Physics	-0.36	(0.363)	-
Subject: Religious studies	-1.902***	(0.176)	-
Subject: Social sciences	-0.071	(0.079)	-
Subject: Swedish	-0.313***	(0.082)	-
Subject: Technology	-0.339	(0.334)	-

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table C.12: Results, random-coefficient multilevel SAOMs. Main model, friend influence only.

	β	SD	varying
Friendship part			
Outdegree (density)	-2.269***	(0.126)	+
Alter's popularity	0.001	(0.046)	+
Ego's popularity	0.117	(0.081)	+
Ego's activity	-0.484***	(0.060)	+
Reciprocity	2.220***	(0.065)	+
Transitive triplets	0.786***	(0.034)	+
Transitive reciprocated triplets	-0.473***	(0.039)	+
Female ego	0.043	(0.032)	-
Female alter	-0.014	(0.027)	-
Same gender of ego and alter	0.498***	(0.026)	-
Shared favourite subject of ego and alter	-0.268***	(0.096)	-
Female ego * shared favourite subject of ego and alter	0.219*	(0.089)	-
Favourite subject part			
Outdegree (density)	-3.136***	(0.114)	+
Subject popularity	0.517***	(0.059)	+
Ego's popularity	-0.073	(0.068)	+
Ego's friendship activity	-0.262***	(0.078)	+
Friend influence	0.984***	(0.073)	-
Female ego * friend influence	-0.375***	(0.065)	-

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table C.13: Results, random-coefficient multilevel SAOMs. Main model, friend influence only, with subject effects.

	β	SD	varying
Friendship part			
Outdegree (density)	-2.247***	(0.141)	+
Alter's popularity	-0.008	(0.050)	+
Ego's popularity	0.142*	(0.082)	+
Ego's activity	-0.502***	(0.063)	+
Reciprocity	2.228***	(0.067)	+
Transitive triplets	0.789***	(0.034)	+
Transitive reciprocated triplets	-0.477***	(0.040)	+
Female ego	0.043	(0.031)	-
Female alter	-0.022	(0.024)	-
Same gender of ego and alter	0.491***	(0.024)	-
Shared favourite subject of ego and alter	-0.295***	(0.091)	-
Female ego * shared favourite subject of ego and alter	0.240***	(0.087)	-
Favourite subject part			
Outdegree (density)	-1.650***	(0.215)	+
Subject popularity	0.083	(0.070)	+
Ego's popularity	-0.044	(0.065)	+
Ego's friendship activity	-0.287***	(0.076)	+
Friend influence	1.013***	(0.070)	-
Female ego * friend influence	-0.382***	(0.061)	-
Parental education ego	-0.025***	(0.008)	-
Parental education ego * STEM subject	-0.034	(0.032)	-
Cognitive ability ego	-0.329***	(0.093)	-
Cognitive ability ego * STEM subject	0.050***	(0.012)	-
Subject: Arts	-1.239***	(0.151)	-
Subject: Biology	-1.355***	(0.330)	-
Subject: Chemistry	-0.265***	(0.092)	-
Subject: Crafts	-0.697***	(0.106)	-
Subject: History	-0.604***	(0.103)	-
Subject: Home economics	0.507*	(0.235)	-
Subject: Maths	-0.243***	(0.086)	-
Subject: Music	-0.171	(0.248)	-
Subject: Natural Sciences	-1.418***	(0.146)	-
Subject: Foreign languages	-0.007	(0.082)	-
Subject: Physical education	-0.782***	(0.257)	-
Subject: Physics	-1.936***	(0.176)	-
Subject: Religious studies	-0.082	(0.079)	-
Subject: Social sciences	-0.294***	(0.098)	-
Subject: Swedish	-0.860***	(0.283)	-
Subject: Technology	-0.481***	(0.098)	-

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table C.14: Results, random-coefficient multilevel SAOMs. Main model, peer exposure only.

	β	SD	varying
Friendship part			
Outdegree (density)	-2.301***	(0.129)	+
Alter's popularity	-0.004	(0.047)	+
Ego's popularity	0.153*	(0.085)	+
Ego's activity	-0.509***	(0.063)	+
Reciprocity	2.225***	(0.066)	+
Transitive triplets	0.791***	(0.034)	+
Transitive reciprocated triplets	-0.484***	(0.039)	+
Female ego	0.076***	(0.027)	-
Female alter	-0.013	(0.026)	-
Same gender of ego and alter	0.487***	(0.024)	-
Favourite subject part			
Outdegree (density)	-3.729***	(0.181)	+
Indegree popularity (sqrt)	0.808***	(0.106)	+
Indegree friendship activity	-0.042	(0.069)	+
Outdegree friendship activity	-0.039	(0.077)	+
Female ego	0.767***	(0.134)	-
Subject popularity among females	0.412***	(0.124)	-
STEM subject	0.280*	(0.175)	-
Female ego * subject popularity among females	-0.436***	(0.078)	-
Female ego * STEM subject	-0.992***	(0.160)	-
STEM subject * subject popularity among females	-0.155	(0.137)	-
Female ego * STEM subject * subject popularity among females	0.440***	(0.125)	-

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table C.15: Results, random-coefficient multilevel SAOMs. Main model, peer exposure only with subject effects.

	β	SD	varying
Friendship part			
Outdegree (density)	-2.305***	(0.132)	+
Alter's popularity	-0.010	(0.049)	+
Ego's popularity	0.132	(0.091)	+
Ego's activity	-0.496***	(0.069)	+
Reciprocity	2.239***	(0.065)	+
Transitive triplets	0.792***	(0.035)	+
Transitive reciprocated triplets	-0.484***	(0.039)	+
Female ego	0.073***	(0.027)	-
Female alter	-0.001	(0.026)	-
Same gender of ego and alter	0.490***	(0.021)	-
Favourite subject part			
Outdegree (density)	-2.517***	(0.261)	+
Indegree popularity (sqrt)	0.511***	(0.129)	+
Indegree friendship activity	-0.040	(0.064)	+
Outdegree friendship activity	-0.024	(0.068)	+
Female ego	0.927***	(0.136)	-
Subject popularity among females	0.472***	(0.137)	-
Female ego * subject popularity among females	-0.533***	(0.090)	-
Female ego * STEM subject	-1.133***	(0.201)	-
STEM subject * subject popularity among females	-0.430***	(0.152)	-
Female ego * STEM subject * subject popularity among females	0.555***	(0.140)	-
Parental education ego	0.019	(0.023)	-
Parental education ego * STEM subject	-0.035	(0.033)	-
Cognitive ability ego	-0.025***	(0.009)	-
Cognitive ability ego * STEM subject	0.047***	(0.013)	-
Subject: Arts	-0.327***	(0.103)	-
Subject: Biology	-1.227***	(0.127)	-
Subject: Chemistry	-0.616	(0.363)	-
Subject: Crafts	-0.227***	(0.089)	-
Subject: History	-0.732***	(0.126)	-
Subject: Home economics	-0.548***	(0.110)	-
Subject: Maths	0.952***	(0.343)	-
Subject: Music	-0.213*	(0.094)	-
Subject: Natural Sciences	0.373	(0.335)	-
Subject: Foreign languages	-1.350***	(0.127)	-
Subject: Physical education	0.141	(0.089)	-
Subject: Physics	-0.103	(0.299)	-
Subject: Religious studies	-1.970***	(0.210)	-
Subject: Social sciences	-0.036	(0.074)	-
Subject: Swedish	-0.277***	(0.093)	-
Subject: Technology	-0.263	(0.313)	-

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table C.16: Proportion of students having a STEM-favourite subjects by having a romantic partner in the classroom and gender

	romantic partner in classroom	wave 1	wave 2
female	yes	14.70% (n = 35)	13.20% (n = 55)
	no	18.67% (n = 2511)	14.82% (n = 2199)
male	yes	21.05% (n = 39)	26.41% (n = 58)
	no	20.52% (n = 2440)	19.01% (n = 2198)

Table C.17: Proportion of students having a STEM-favourite subjects by having a romantic partner and gender

	romantic partner	wave 1	wave 2
female	yes	19.14% (n = 426)	12.78% (n = 502)
	no	18.51% (n = 2120)	15.35% (n = 1752)
male	yes	19.09% (n = 343)	18.58% (n = 372)
	no	20.76% (n = 2136)	19.31% (n = 1884)

Appendix D

Appendix Chapter 5

Table D.1: Variable descriptives: analytical and full sample.

	analytical sample						*full sample*					
	mean	SD	median	min	max	N	mean	SD	median	min	max	N
Maths competence beliefs, wave 1, low to high	3.74	0.97	4.00	1.00	5.00	4815	3.74	0.96	4.00	1.00	5.00	4996
Maths competence beliefs, wave 2, low to high	3.64	1.04	4.00	1.00	5.00	4396	3.64	1.04	4.00	1.00	5.00	4513
Grades in maths, low to high	2.49	0.81	2.00	1.00	4.00	4368	2.49	0.81	2.00	1.00	4.00	4484
Gender: female						2647						2744
Gender: male						2583						2683

Table D.2: Friendship network descriptives, included and dropped groups.

	included groups				*dropped groups*			
	min	max	mean	SD	min	max	mean	SD
size	10.00	34.00	22.11	4.32	6.00	28.00	14.93	7.56
density	0.01	0.43	0.13	0.04	0.00	0.47	0.15	0.12
degrees	0.00	5.00	2.71	1.85	0.00	5.00	1.70	1.85
avg degree	0.12	4.70	2.66	0.76	0.00	3.44	1.63	1.20
clustering	0.00	1.00	0.57	0.12	0.00	1.00	0.64	0.28
jaccard	0.04	0.58	0.32	0.10	0.00	0.83	0.27	0.32
same gender ties (t1)	0.54	1.00	0.90	0.08	0.88	1.00	0.96	0.04
same gender ties (t2)	0.51	1.00	0.87	0.10	0.86	1.00	0.97	0.04

Table D.3: Details on model specification (Selection part of the model).

Effect name (*internal RSiena effect name*)	The tendency is modelled...
Selection part. *Dependent variable: friendship (ego creates or maintains a tie to alter)*	
Outdegree (*density*)	... to create and maintain friendships.
Reciprocity (*recip*)	... to reciprocate friendships.
Transitive triplets (*transTrip*)	... to create and maintain friendships to friends of friends.
Transitive reciprocated triplets (*transRecTrip*)	... to reciprocate ties to friends of friends.
Alter's popularity (*inPop*)	... to create and maintain ties to actors that are popular in the network.
Ego's popularity (*inAct*)	... to create and maintain ties to actors that have ties to more others.
Ego's activity (*outAct*)	... for actors who already have more ties, to create and maintain more
Female alter (*altX female*)	... to create and maintain more ties to girls than to boys
Female ego (*egoX female*)	... for girls to create and maintain more ties
Same gender of ego and alter (*sameX female*)	... for actors to be tied to those of the same gender of ego and alter
ASC alter (*altX ASC*)	... to create and maintain more ties to actors with higher levels of maths competence beliefs
ASC ego (*egoX ASC*)	... for actors with higher levels of maths competence beliefs to create and maintain more ties
Friend selection (*simX ASC*)	... for actors to be tied to those with similar levels of maths competence beliefs

Table D.4: Details on model specification (Influence part of the model).

Effect name (*internal RSiena effect name*)	The tendency is modelled...
Influence part: *Dependent variable: maths competence beliefs (ego increases or maintains his or her maths competence beliefs)*	
Linear shape (*linear*)	... for maths competence beliefs to increase
Quadratic shape (*quadtratic*)	... for amaths competence beliefs to increase exponentially
Effect from female (*effFrom female*)	... for girls to have higher levels of maths competence beliefs
Effect from maths grades (*effFrom maths grades*)	... for actors with higher maths grades to have higher maths competence beliefs
Effect from maths grades * Effect from female (*effFrom maths grades * effFrom female*)	... for the effect from maths grades on maths competence beliefs to apply differently to girls
Friend influence (*avSim MCB*)	... for actors to change their maths competence beliefs towards the average maths competence beliefs of his or her friends
Friend influence * Effect from female (*avSim MCB * effFrom female*)	... for friend influence to apply differently to girls
Friend influence increase (*creation: avSim MCB*)	... for actors to increase their maths competence beliefs towards the average maths competence beliefs of his or her friends
Friend influence decrease (*endowment: avSim MCB*)	... for actors to decrease their maths competence beliefs towards the average maths competence beliefs of his or her friends
Friend influence increase * Effect from female (*creation: avSim MCB * effFrom female*)	... for friend influence increase to apply differently to girls
Friend influence decrease * Effect from female (*endowment: avSim MCB * effFrom female*)	... for friend influence decrease to apply differently to girls

Table D.5: Model 1a, full results. Random coefficient multilevel SAOMs, friendship selection and influence on maths confidence beliefs.

	β	SD	varying
Selection part			
Outdegree	-2.424***	(0.082)	+
Reciprocity	2.319***	(0.063)	+
Transitive triplets	0.859***	(0.032)	+
Transitive reciprocated triplets	-0.481***	(0.041)	+
Alter's popularity	-0.042*	(0.020)	+
Ego's activity	-0.154***	(0.024)	+
Ego's popularity	-0.009	(0.044)	+
Female alter	0.000	(0.027)	+
Female ego	0.055*	(0.024)	+
Same gender of ego and alter	0.451***	(0.025)	+
ASC alter	-0.011	(0.011)	-
ASC ego	0.096***	(0.013)	-
Friend selection	0.106*	(0.052)	-
Influence part			
Linear shape	0.098*	(0.051)	+
Quadratic shape	-0.028^{+}	(0.022)	+
Effect from female	-0.242***	(0.045)	-
Friend influence	0.854***	(0.228)	-

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{+}p < 0.1$

Table D.6: Model 1b, full results. Random coefficient multilevel SAOMs, friendship selection and influence on maths confidence belief.

	β	SD	varying
Selection part			
Outdegree	-2.659***	(0.101)	+
Reciprocity	2.380***	(0.069)	+
Transitive triplets	0.881***	(0.036)	+
Transitive reciprocated triplets	-0.492***	(0.043)	+
Alter's popularity	-0.044*	(0.023)	+
Ego's activity	-0.170***	(0.025)	+
Ego's popularity	0.017	(0.042)	+
Female alter	-0.010	(0.051)	+
Female ego	0.080^{+}	(0.054)	+
Same gender of ego and alter	0.584***	(0.052)	+
ASC alter	-0.009	(0.014)	-
ASC ego	0.099***	(0.015)	-
Friend selection	0.073	(0.058)	-
Influence part			
Linear shape	0.344***	(0.058)	+
Quadratic shape	-0.397***	(0.032)	+
Effect from female	-0.442***	(0.053)	-
Friend influence	0.397^{+}	(0.296)	-
Effect from math grades	1.195***	(0.105)	+
Effect from maths grades * effect from female	0.236***	(0.083)	+

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{+}p < 0.1$

Table D.7: Model 2a, full results. Random coefficient multilevel SAOMs, friendship selection and influence on maths confidence belief.

	β	SD	varying
Selection part			
Outdegree	-2.671***	(0.102)	+
Reciprocity	2.359***	(0.070)	+
Transitive triplets	0.871***	(0.036)	+
Transitive reciprocated triplets	-0.482***	(0.043)	+
Alter's popularity	-0.041*	(0.023)	+
Ego's activity	-0.169***	(0.024)	+
Ego's popularity	0.012	(0.039)	+
Female alter	-0.016	(0.050)	+
Female ego	0.100*	(0.057)	+
Same gender of ego and alter	0.589***	(0.049)	+
ASC alter	-0.002	(0.013)	-
ASC ego	0.112***	(0.015)	-
Friend selection	0.103*	(0.056)	-
Influence part			
Linear shape	0.067	(0.050)	+
Quadratic shape	-0.025	(0.022)	+
Effect from female	-0.220***	(0.045)	-
Friend influence	1.635***	(0.380)	-
Friend influence * effect from female	-0.703*	(0.293)	-

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{+}p < 0.1$

Table D.8: Model 2b, full results. Random coefficient multilevel SAOMs, friendship selection and influence on maths confidence belief.

	β	SD	varying
Selection part			
Outdegree	-2.665***	(0.107)	+
Reciprocity	2.374***	(0.068)	+
Transitive triplets	0.876***	(0.036)	+
Transitive reciprocated triplets	-0.488***	(0.042)	+
Alter's popularity	-0.043*	(0.023)	+
Ego's activity	-0.170***	(0.024)	+
Ego's popularity	0.014	(0.040)	+
Female alter	-0.013	(0.051)	+
Female ego	0.084	(0.054)	+
Same gender of ego and alter	0.591***	(0.052)	+
ASC alter	-0.007	(0.013)	-
ASC ego	0.103***	(0.015)	-
Friend selection	0.065	(0.055)	-
Influence part			
Linear shape	0.338***	(0.062)	+
Quadratic shape	-0.395***	(0.033)	+
Effect from female	-0.436***	(0.058)	+
Friend influence	0.508	(0.492)	-
Friend influence * effect from female	-0.053	(0.386)	-
Effect from math grades	1.210***	(0.114)	+
Effect from math grades * effect from female	0.224***	(0.089)	+

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{+}p < 0.1$

Table D.9: Model 3a, full results. Random coefficient multilevel SAOMs, friendship selection and influence on maths confidence belief.

	β	SD	varying
Selection part			
Outdegree	-2.671***	(0.105)	+
Reciprocity	2.374***	(0.070)	+
Transitive triplets	0.873***	(0.037)	+
Transitive reciprocated triplets	-0.486***	(0.046)	+
Alter's popularity	-0.039*	(0.022)	+
Ego's activity	-0.167***	(0.025)	+
Ego's popularity	0.012	(0.040)	+
Female alter	-0.013	(0.052)	+
Female ego	0.094*	(0.056)	+
Same gender of ego and alter	0.583***	(0.049)	+
ASC alter	0.003	(0.016)	-
ASC ego	0.113***	(0.017)	-
Friend selection	0.098^{+}	(0.062)	-
Influence part			
Linear shape	0.170***	(0.064)	+
Quadratic shape	-0.013	(0.022)	+
Effect from female	-0.232***	(0.054)	-
Friend influence (decrease)	-0.081	(0.627)	-
Friend influence (increase)	3.829***	(0.789)	-
Friend influence (decrease) * effect from female	-0.416	(0.533)	-
Friend influence (increase) * effect from female	-1.016^{+}	(0.651)	-

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{+}p < 0.1$

Table D.10: Model 3b, full results. Random coefficient multilevel SAOMs, friendship selection and influence on maths confidence belief.

	β	SD	varying
Selection part			
Outdegree	-2.668***	(0.110)	+
Reciprocity	2.382***	(0.070)	+
Transitive triplets	0.881***	(0.037)	+
Transitive reciprocated triplets	-0.494***	(0.044)	+
Alter's popularity	-0.043*	(0.023)	+
Ego's activity	-0.174***	(0.026)	+
Ego's popularity	0.024	(0.041)	+
Female alter	-0.015	(0.049)	+
Female ego	0.084+	(0.056)	+
Same gender of ego and alter	0.587***	(0.051)	+
ASC alter	-0.006	(0.016)	-
ASC ego	0.105***	(0.015)	-
Friend selection	0.055	(0.063)	-
Influence part			
Linear shape	0.393***	(0.066)	+
Quadratic shape	-0.380***	(0.032)	+
Effect from female	-0.437***	(0.068)	-
Friend influence (decrease)	-0.369	(0.762)	-
Friend influence (increase)	1.800*	(0.888)	-
Friend influence (decrease) * effect from female	-0.213	(0.635)	-
Friend influence (increase) * effect from female	0.267	(0.755)	-
Effect from math grades	1.198***	(0.119)	+
Effect from maths grades * effect from female	0.235*	(0.101)	+

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{+}p < 0.1$

Table D.11: Random coefficient multilevel SAOMs, friendship selection and influence on maths confidence belief.

	β	SD	varying
Selection part			
Outdegree	-2.449***	(0.079)	+
Reciprocity	2.335***	(0.071)	+
Transitive triplets	0.856***	(0.033)	+
Transitive reciprocated triplets	-0.473***	(0.040)	+
Alter's popularity	-0.040*	(0.020)	+
Ego's activity	-0.153***	(0.022)	+
Ego's popularity	-0.016	(0.038)	+
Female alter	0.005	(0.029)	+
Female ego	0.060*	(0.029)	+
Same gender of ego and alter	0.451***	(0.026)	+
ASC alter	-0.007	(0.013)	-
ASC ego	0.098***	(0.015)	-
Friend selection	0.105^{+}	(0.063)	-
Influence part			
Linear shape	0.154***	(0.054)	+
Quadratic shape	-0.009	(0.027)	+
Effect from female	-0.215***	(0.045)	-
Friend influence (decrease)	0.148	(0.452)	-
Friend influence (increase)	3.472***	(0.565)	-
Friend influence * effect from female	-0.664*	(0.299)	-

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{+}p < 0.1$

Table D.12: Random coefficient multilevel SAOMs, friendship selection and influence on maths confidence belief.

	β	SD	varying
Selection part			
Outdegree	-2.673***	(0.105)	+
Reciprocity	2.374***	(0.070)	+
Transitive triplets	0.880***	(0.036)	+
Transitive reciprocated triplets	-0.490***	(0.043)	+
Alter's popularity	-0.043*	(0.024)	+
Ego's activity	-0.174***	(0.025)	+
Ego's popularity	0.020	(0.040)	+
Female alter	-0.015	(0.051)	+
Female ego	0.091+	(0.057)	+
Same gender of ego and alter	0.598***	(0.049)	+
ASC alter	-0.004	(0.015)	-
ASC ego	0.104***	(0.016)	-
Friend selection	0.076+	(0.061)	-
Influence part			
Linear shape	0.387***	(0.064)	+
Quadratic shape	-0.382***	(0.033)	+
Effect from female	-0.436***	(0.055)	-
Friend influence (decrease)	-0.565	(0.570)	-
Friend influence (increase)	2.054***	(0.598)	-
Friend influence * effect from female	-0.034	(0.393)	-
Effect from math grades	1.178***	(0.121)	+
Effect from maths grades * effect from female	0.255*	(0.103)	+

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{+}p < 0.1$

Table D.13: Random coefficient multilevel SAOMs, friendship selection and influence on maths confidence belief.

	β	SD	varying
Selection part			
Outdegree	-2.443***	(0.081)	+
Reciprocity	2.330***	(0.068)	+
Transitive triplets	0.856***	(0.033)	+
Transitive reciprocated triplets	-0.475***	(0.040)	+
Alter's popularity	-0.040*	(0.020)	+
Ego's activity	-0.152***	(0.021)	+
Ego's popularity	-0.019	(0.035)	+
Female alter	0.004	(0.026)	+
Female ego	0.065*	(0.028)	+
Same gender of ego and alter	0.447***	(0.025)	+
ASC alter	-0.009	(0.012)	-
ASC ego	0.102***	(0.016)	-
Friend selection	0.090^{+}	(0.054)	-
Influence part			
Linear shape	0.071^{+}	(0.052)	+
Quadratic shape	-0.011	(0.023)	+
Effect from female	-0.160***	(0.048)	-
Friend influence	1.707***	(0.414)	-
Friend influence (decrease) * effect from female	-1.769***	(0.394)	-
Friend influence (increase) * effect from female	0.615^{+}	(0.441)	-

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{+}p < 0.1$

Table D.14: Random coefficient multilevel SAOMs, friendship selection and influence on maths confidence belief.

	β	SD	varying
Selection part			
Outdegree	-2.451***	(0.089)	+
Reciprocity	2.337***	(0.067)	+
Transitive triplets	0.856***	(0.033)	+
Transitive reciprocated triplets	-0.475***	(0.039)	+
Alter's popularity	-0.039*	(0.021)	+
Ego's activity	-0.151***	(0.021)	+
Ego's popularity	-0.019	(0.036)	+
Female alter	0.007	(0.027)	+
Female ego	0.054*	(0.031)	+
Same gender of ego and alter	0.448***	(0.025)	+
ASC alter	-0.013	(0.013)	-
ASC ego	0.091***	(0.015)	-
Friend selection	0.061	(0.061)	-
Influence part			
Linear shape	0.348***	(0.061)	+
Quadratic shape	-0.378***	(0.030)	+
Effect from female	-0.397***	(0.056)	-
Friend influence	0.959*	(0.488)	-
Friend influence (decrease) * effect from female	-1.309***	(0.438)	-
Friend influence (increase) * effect from female	0.716^{+}	(0.477)	-
Effect from math grades	1.221***	(0.075)	+
Effect from maths grades * effect from female	0.215*	(0.145)	+

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{+}p < 0.1$

Bibliography

Åberg, Y. and Hedström, P. (2001). Youth unemployment: A self-reinforcing process? Cambridge University Press, Cambridge.

Agirdag, O., Van Houtte, M., and Van Avermaet, P. (2012). Why Does the Ethnic and Socio-economic Composition of Schools Influence Math Achievement? The Role of Sense of Futility and Futility Culture. *European Sociological Review*, 28(3):366–378.

Akerlof, G. A. and Kranton, R. E. (2000). Economics and Identity. *The Quarterly Journal of Economics*, 115(3):715–753.

Allison, P. D. (2009). *Fixed effects regression models*. Sage university papers. Quantitative applications in the social sciences. SAGE, London.

Alper, J. (1993). The Pipeline Is Leaking Women All the Way Along. *Science*, 260(5106):409–411.

Arens, A. K., Yeung, A. S., Craven, R. G., and Hasselhorn, M. (2011). The twofold multidimensionality of academic self-concept: Domain specificity and separation between competence and affect components. *Journal of Educational Psychology*, 103(4):970–981.

Babcock, P. (2008). From Ties to Gains? Evidence on Connectedness and Human Capital Acquisition. *Journal of Human Capital*, 2(4):379–409.

Bandura, A. (1977). *Social learning theory*. Prentice-Hall, Englewood Cliffs.

Barry, C. M. and Wentzel, K. R. (2006). Friend influence on prosocial behavior: The

role of motivational factors and friendship characteristics. *Developmental Psychology*, 42(1):153–163.

Baumeister, R. F. (2011). Self and identity. *Annals of the New York Academy of Sciences*, 1234(1):48–55.

Baumeister, R. F. and Leary, M. R. (1995). The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, 117(3):497–529.

Bicchieri, C. (2006). *The grammar of society : the nature and dynamics of social norms*. Cambridge University Press, Cambridge.

Blau, P. M. and Duncan, O. D. (1967). *The American Occupational Structure*. Wiley, New York.

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

Block, P., Heathcote, L. C., and Burnett Heyes, S. (2018). Social interaction and pain. *Social Science & Medicine*, 196:47–55.

Bobba, M. and Gignoux, J. (2011). Policy-induced Social Interactions and Schooling Decisions.

Boda, Z. (2018). Social Influence on Observed Race. *Sociological Science*, 5:29–57.

Borsari, B. and Carey, K. B. (2001). Peer influences on college drinking: A review of the research. *Journal of substance abuse*, 13(4):391–424.

Breen, R., editor (2004). *Social Mobility in Europe*. Oxford University Press, Oxford.

Breen, R. and Jonsson, J. O. (2005). Inequality of Opportunity in Comparative Perspective: Recent Research on Educational Attainment and Social Mobility. *Annual Review of Sociology*, 31(1):223–243.

Brinbaum, Y. and Cebolla-Boado, H. (2007). The school careers of ethnic minority youth in France: Success or disillusion? *Ethnicities*, 7(3):445–474.

Brown, B. B. (2004). *Handbook of Adolescent Psychology*. Wiley, Hoboken, 2nd edition.

Brown, B. B., Clasen, D. R., and Eicher, S. A. (1986). Perceptions of peer pressure, peer conformity dispositions, and self-reported behavior among adolescents. *Developmental Psychology*, 22(4):521–530.

Browne, J. (2006). *Sex segregation and inequality in the modern labour market*. Ebook central. Policy Press, Bristol, UK.

Burgess, R. L. and Akers, R. L. (1966). A differential association-reinforcement theory of criminal behavior. *American Sociological Review*, 14(2):128–147.

Burt, R. S. (2007). Structural Holes versus Network Closure as Social Capital. In Lin, N., Cook, K. S., and Burt, R. S., editors, *Social Capital*.

Byrne, D. (1997). An overview (and underview) of research and theory within the attraction paradigm. *Journal of Social and Personal Relationships*, 14(3):417–431.

Carbonaro, W. and Workman, J. (2016). Intermediate peer contexts and educational outcomes: Do the friends of students' friends matter? *Social Science Research*, 58:184–197.

Carrington, P. J., Scott, J., and Wasserman, S., editors (2009). *Models and Methods in Social Network Analysis*. 27. Cambridge University Press, Cambridge.

Charles, M. and Bradley, K. (2002). Equal but separate? *American Sociological Review*, 67(4):573.

Charles, M. and Grusky, D. B. (2004). *Occupational ghettos*. The worldwide segregation of women and men. Stanford University Press, Stanford, California.

Cheadle, J. E., Walsemann, K. M., and Goosby, B. J. (2015). Teen Alcohol Use and Social Networks: The Contributions of Friend Influence and Friendship Selection. *Journal of Alcoholism & Drug Dependence*, 3(5):1–9.

Chen, F. R., Rothman, E. F., and Jaffee, S. R. (2017). Early Puberty, Friendship Group Characteristics, and Dating Abuse in US Girls. *Pediatrics*, 139(6):1–8.

Chow, A. and Salmela-Aro, K. (2011). Task-values across subject domains: A gender comparison using a person-centered approach. *International Journal of Behavioral Development*, 35(3):202–209.

CILS4EU (2014). Children of Immigrants Longitudinal Survey in Four European Countries. Technical Report. Wave 1 – 2010/2011, v1.1.0. . Technical report, Mannheim.

Coleman, J. S. (1987). Microfoundations and macrosocial behavior. In *The Micro-macro link*, pages 153–173. University of California Press, Berkeley.

Coleman, J. S. (1988). Social Capital in the Creation of Human Capital. *American Journal of Sociology*, 94.

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

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., and York, R. L. (1966). Equality of Educational Opportunity.

Corak, M. (2013). Income Inequality, Equality of Opportunity, and Intergenerational Mobility. *Journal of Economic Perspectives*, 27(3):79–102.

Correll, S. J. (2001). Gender and the Career Choice Process: The Role of Biased Self-Assessments. *American Journal of Sociology*, 106(6):1691–1730.

Correll, S. J. (2004). Constraints into Preferences: Gender, Status, and Emerging Career Aspirations. *American Sociological Review*, 69(1):93–113.

Crombie, G., Sinclair, N., Silverthorn, N., Byrne, B. M., DuBois, D. L., and Trinneer, A. (2005). Predictors of Young Adolescents? Math Grades and Course Enrollment Intentions: Gender Similarities and Differences. *Sex Roles*, 52(5-6):351–367.

Crul, M. and Vermeulen, H. (2003). The Second Generation in Europe. *The International Migration Review*, 37(4):965–986.

Davis, J. A. (1970). Clustering and Hierarchy in Interpersonal Relations: Testing Two Graph Theoretical Models on 742 Sociomatrices. *American Sociological Review*, 35(5):843–851.

DeLay, D., Zhang, L., Hanish, L. D., Miller, C. F., Fabes, R. A., Martin, C. L., Kochel, K. P., and Updegraff, K. A. (2016). Peer Influence on Academic Performance: A Social Network Analysis of Social-Emotional Intervention Effects. *Prevention Science*, 17:903–913.

Dishion, T. J. and Tipsord, J. M. (2011). Peer Contagion in Child and Adolescent Social and Emotional Development. *Annual review of psychology*, 62(1):189–214.

Dollmann, J., Jacob, K., and Kalter, F. (2014). Examining the Diversity of Youth in Europe. Technical report, Mannheim.

Dryler, H. (1999). The impact of school and classroom characteristics on educational choices by boys and girls: A multilevel analysis. *Acta Sociologica*, 42(4):299–318.

Dustmann, C. and Frattini, T. (2012). Immigration: The European Experience. *NORFACE Migration Discussion Paper*, 01.

Duvander, A.-Z. and Andersson, G. (2006). Gender Equality and Fertility in Sweden. *Marriage & Family Review*, 39(1-2):121–142.

Efferson, C., Roca, C. P., Vogt, S., and Helbing, D. (2016). Sustained cooperation by running away from bad behavior. *Evolution and Human Behavior*, 37(1):1–9.

Eitle, D., Thorsen, M. L., and McNulty Eitle, T. (2016). School Context, Race, and Adolescent Risky Sexual Behaviors. *European Conference on Social and Behavioral Sciences, International Association of Social Science Research.*

Elmer, T., Boda, Z., and Stadtfeld, C. (2017). The co-evolution of emotional well-being with weak and strong friendship ties. *Network Science*, 3(5):1–34.

England, P. (2010). The Gender Revolution. *Gender & Society*, 24(2):149–166.

England, P., Allison, P., Li, S., Mark, N., Thompson, J., Budig, M. J., and Sun, H. (2007). Why are some academic fields tipping toward female? The sex composition of US fields of doctoral degree receipt, 1971–2002. *Acta Sociologica*, 80(1):23–42.

Erikson, R. and Goldthorpe, J. H. (1992). *The constant flux : a study of class mobility in industrial societies*. Clarendon Press, Oxford.

Erikson, R. and Jonsson, J. O., editors (1996). *Can education be equalized? The Swedish case in comparative perspective*. Westview Press, Boulder.

Ermisch, J. (2008). Origins of Social Immobility and Inequality: Parenting and Early Child Development. *National Institute Economic Review*, 205:62–71.

EuropeanCommission (2016). She Figures 2015. European Union.

Everett, M. G. and Borgatti, S. P. (1998). Analyzing clique overlap. *Connections*.

Feinstein, L. (2003). Inequality in the Early Cognitive Development of British Children in the 1970 Cohort. *Economica*, 70:73–97.

Feliciano, C. and Rumbaut, R. G. (2005). Gendered paths: Educational and occupational expectations and outcomes among adult children of immigrants. *Ethnic and Racial Studies*, 28(6):1087–1118.

Fleischmann, F. and Dronkers, J. (2010). Unemployment among immigrants in European labour markets: an analysis of origin and destination effects. *Work, Employment and Society*, 24(2):337–354.

Fordham, S. and Ogbu, J. U. (1986). Black Students', School Success: Coping With the "Burden of 'Acting White"'. *Urban Review*.

Friedkin, N. E. (2001). Norm formation in social influence networks. *Social Networks*, 23(3):167–189.

Fryer, R. G. and Torelli, P. (2010). An empirical analysis of 'acting white'. *Journal of Public Economics*, 94(5):380–396.

Gabay-Egozi, L., Shavit, Y., and Yaish, M. (2014). Gender Differences in Fields of Study: The Role of Significant Others and Rational Choice Motivations. *European Sociological Review*, 31(3):284–297.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013). *Bayesian data analysis*. Texts in statistical science series. CRC Press, Boca Raton, Florida.

Gerber, T. P. and Cheung, S. Y. (2008). Horizontal Stratification in Postsecondary Education: Forms, Explanations, and Implications. *Annual Review of Sociology*, 34(1):299–318.

Geven, S., Jonsson, J. O., and Tubergen, F. (2017). Gender Differences in Resistance to Schooling: The Role of Dynamic Peer-Influence and Selection Processes. *Journal of Youth and Adolescence*, 46:2421–2445.

Granovetter, M. (1978). Threshold Models of Collective Behavior. *American Journal of Sociology*, 83(6):1420–1443.

Granovetter, M. S. (1973). The Strength of Weak Ties. *American Journal of Sociology*, 78(6):1360–1380.

Haas, S. A. and Schaefer, D. R. (2014). With a Little Help from My Friends? Asymmetrical Social Influence on Adolescent Smoking Initiation and Cessation. *Journal of Health and Social Behavior*, 55(2):126–143.

Hauser, R. M., Tsai, S.-L., and Sewell, W. H. (1983). A model of stratification with response error in social and psychological variables. *Sociology of Education*, 56(1):20.

Heath, A. F. (2007). Crossnational patterns and processes of ethnic disadvantage. In *Unequal chances : ethnic minorities in Western labour markets*. Oxford University Press, Oxford.

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

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

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*.

Hjalmarsson, S. and Mood, C. (2015). Do poorer youth have fewer friends? The role of household and child economic resources in adolescent school-class friendships. *Children and Youth Services Review*, 57:201–211.

Hornsey, M. J. (2008). Social Identity Theory and Self-categorization Theory: A Historical Review. *Social and Personality Psychology Compass*, 2(1):204–222.

Huang, C. (2011). Self-concept and academic achievement: A meta-analysis of longitudinal relations. *Journal of School Psychology*, 49(5):505–528.

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, 60(6):581–592.

Ichou, M. (2014). Who They Were There: Immigrants' Educational Selectivity and Their Children's Educational Attainment. *European Sociological Review*, 30(6):750–765.

Inglehart, R. and Norris, P. (2003). *Rising tide : gender equality and cultural change around the world*. Cambridge University Press, Cambridge.

Inglehart, R. F. (2008). Changing Values among Western Publics from 1970 to 2006. *West European Politics*, 31(1-2):130–146.

Ioannides, Y. M. and Datcher Loury, L. (2004). Job Information Networks, Neighborhood Effects.

Jackson, M., editor (2013). *Determined to Succeed?* Performance versus Choice in Educational Attainment. Stanford University Press, Stanford, California.

Jackson, M., Jonsson, J. O., and Rudolphi, F. (2012). Ethnic Inequality in Choice-driven Education Systems: A Longitudinal Study of Performance and Choice in England and Sweden. *Sociology of Education*, 85(2):158–178.

Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., and Wigfield, A. (2002). Changes in Children's Self-Competence and Values: Gender and Domain Differences across Grades One through Twelve. *Child Development*, 73(2):509–527.

Jencks, C., Crouse, J., and Mueser, P. (1983). The Wisconsin model of status attainment. *Sociology of Education*, 56(1):3.

Jonsson, J. O. (1999). Explaining Sex Differences in Educational Choice: An Empirical Assessment of a Rational Choice Model. *European Sociological Review*, 15(4):391–404.

Jonsson, J. O. and Mood, C. (2008). Choice by contrast: How peers' achievement affects educational choice in Sweden. *Social Forces*, 87:741–765.

Jonsson, J. O. and Rudolphi, F. (2011). Weak Performance–Strong Determination: School Achievement and Educational Choice among Children of Immigrants in Sweden. *European Sociological Review*, 27(4):487–508.

Kalter, F., Heath, A., Hewstone, M., Jonsson, J. O., Kalmijn, M., Kogan, I., and Tubergen, F. v. (2014). Children of Immigrants Longitudinal Survey in Four European Countries (CILS4EU) .

Kalter, F., Heath, A., Hewstone, M., Jonsson, J. O., Kalmijn, M., Kogan, I., and Tubergen, F. v. (2015). Children of Immigrants Longitudinal Survey in Four European Countries (CILS4EU).

Kobus, K. (2003). Peers and adolescent smoking. *Addiction*, 98(1):37–55.

Köller, O., Trautwein, U., Lüdtke, O., and Baumert, J. (2006). Zum Zusammenspiel von schulischer Leistung, Selbstkonzept und Interesse in der gymnasialen Oberstufe. *Zeitschrift für Pädagogische Psychologie*, 20(1/2):27–39.

Kossinets, G. and Watts, D. J. (2009). Origins of Homophily in an Evolving Social Network. *American Journal of Sociology*, 115(2):405–450.

Kroska, A. (2007). *Gender Ideology and Gender Role Ideology*, volume 60. John Wiley & Sons, Ltd, Oxford, UK.

Kruse, H. and Jacob, K. (2014). Sociometric Fieldwork Report Children of Immigrants Longitudinal Survey in 4 European Countries. pages 1–11.

Lakon, C. M., Wang, C., Butts, C. T., Jose, R., and Hipp, J. R. (2017). Cascades of emotional support in friendship networks and adolescent smoking. *PLOS ONE*, 12(6):1–20.

Legewie, J. and DiPrete, T. A. (2012). School Context and the Gender Gap in Educational Achievement. *American Sociological Review*, 77(3):463–485.

Legewie, J. and DiPrete, T. A. (2014). The High School Environment and the Gender Gap in Science and Engineering. *Sociology of Education*, 87(4):259–280.

Levanon, A. and Grusky, D. B. (2016). The Persistence of Extreme Gender Segregation in the Twenty-first Century. *American Journal of Sociology*, 122(2):573–619.

Lin, N. (1999). Building a Network Theory of Social Capital. *Connections*, 22(1):28–51.

Lin, N. (2000). Inequality in Social Capital. *Contemporary Sociology*, 29(6):785–795.

Lomi, A., Snijders, T. A. B., Steglich, C. E. G., and Torló, V. J. (2011). Why are some more peer than others? Evidence from a longitudinal study of social networks and individual academic performance. *Social Science Research*, 40:1506–1520.

Luce, R. D. and Perry, A. D. (1949). A method of matrix analysis of group structure. *Psychometrika*, 14(2):95–116.

Marsh, H. W. and Craven, R. G. (2006). Reciprocal Effects of Self-Concept and Performance from a Multidimensional Perspective: Beyond Seductive Pleasure and Unidimensional Perspectives. *Perspectives on Psychological Science*, 1(2):133–163.

Marsh, H. W., Köller, O., and Baumert, J. (2001). Reunification of East and West German School Systems: Longitudinal Multilevel Modelling Study of the Big-Fish-Little-Pond Effect on Academic Self-Concept.

Marsh, H. W. and Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology*, 81(1):59–77.

Marsh, H. W. and Shavelson, R. (1985). Self-Concept: Its Multifaceted, Hierarchical Structure. *Educational Psychologist*, 20(3):107–123.

Marsh, H. W. and Yeung, A. S. (1997). Coursework selection: Relations to academic self-concept and achievement. *American Educational Research Journal*, 34(4):691–720.

Martin, C. L. and Fabes, R. A. (2001). The stability and consequences of young children's same-sex peer interactions. *Developmental Psychology*, 37(3):431–446.

McDaniel, A. (2016). The Role of Cultural Contexts in Explaining Cross-National Gender Gaps in STEM Expectations. *European Sociological Review*, 32(1):122–133.

McDougall, P. and Vaillancourt, T. (2015). Long-term adult outcomes of peer victimization in childhood and adolescence: Pathways to adjustment and maladjustment. *American Psychologist*, 70(4):300–310.

McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, 27:415–444.

Merton, R. K. (1968). The Matthew Effect in Science. *Science*, 159(3810):56–63.

Möller, J., Retelsdorf, J., Köller, O., and Marsh, H. W. (2011). The Reciprocal Internal/External Frame of Reference Model. *American Educational Research Journal*, 48(6):1315–1346.

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

Musterd, S. (2005). Social and Ethnic Segregation in Europe. *Journal of Urban Affairs*, 27(3):331–348.

Nakamoto, J. and Schwartz, D. (2010). Is Peer Victimization Associated with Academic Achievement? A Meta-analytic Review. *Social Development*, 19(2):221–242.

Niepel, C., Brunner, M., and Preckel, F. (2014). Achievement goals, academic self-concept, and school grades in mathematics: Longitudinal reciprocal relations in above

average ability secondary school students. *Contemporary Educational Psychology*, 39(4):301–313.

Osterman, K. F. (2000). Students' need for belonging in the school community. *Review of Educational Research*.

Plenty, S. and Jonsson, J. O. (2017). Social Exclusion among Peers: The Role of Immigrant Status and Classroom Immigrant Density. *Journal of Youth and Adolescence*, 46(6):1275–1288.

Portes, A., Aparicio, R., Haller, W., and Vickstrom, E. (2010). Moving Ahead in Madrid: Aspirations and Expectations in the Spanish Second Generation. *The International Migration Review*, 44(4):767–801.

Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. Simon & Schuster, London.

Rayner, K. E., Schniering, C. A., Rapee, R. M., Taylor, A., and Hutchinson, D. M. (2013). Adolescent girls' friendship networks, body dissatisfaction, and disordered eating: Examining selection and socialization processes. *Journal of Abnormal Psychology*, 122(1):93–104.

Reskin, B. F., Roos, P. A., Donato, K. M., Phipps, P. A., Thomas, B. J., Bird, C. E., Detman, L. A., and Steiger, T. (1990). *Job Queues, Gender Queues*. Temple University Press.

Ripley, R. M., Snijders, T. A. B., Boda, Z., Vörös, A., and Preciado, P. (2018). *Manual for SIENA version 4.0 (version February 16, 2018)* . University of Oxford, Department of Statistics; Nuffield College. http://www.stats.ox.ac.uk/siena/, Oxford.

Rjosk, C., Richter, D., Hochweber, J., Lüdtke, O., Klieme, E., and Stanat, P. (2014). Socioeconomic and language minority classroom composition and individual reading achievement: The mediating role of instructional quality. *Learning and Instruction*, 32:63–72.

Rose, A. J. and Rudolph, K. D. (2006). A review of sex differences in peer relationship processes: Potential trade-offs for the emotional and behavioral development of girls and boys. *Psychological Bulletin*, 132(1):98–131.

Schaefer, D. R. (2016). A Network Analysis of Factors Leading Adolescents to Befriend Substance-Using Peers. *Journal of Quantitative Criminology*, pages 1–38.

Schelling, T. C. (1971). Dynamic models of segregation. *The Journal of Mathematical Sociology*, 1(2):143–186.

Sewell, W. H., Haller, A. O., and Portes, A. (1969). The educational and early occupational attainment process. *American Sociological Review*, 34(1):82.

Shavelson, R. J., Hubner, J. J., and Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Acta Sociologica*, 46(3):407–441.

Sikora, J. and Pokropek, A. (2012). Gender segregation of adolescent science career plans in 50 countries. *Science Education*, 96(2):234–264.

Sinclair, S. and Carlsson, R. (2013). What will I be when I grow up? The impact of gender identity threat on adolescents' occupational preferences. *Journal of Adolescence*, 36(3):465–474.

Smith, E. (2011). Women into science and engineering? Gendered participation in higher education STEM subjects. *British Educational Research Journal*, 37(6):993–1014.

Snijders, T. A. B. (2001). The Statistical Evaluation of Social Network Dynamics. *Sociological Methodology*, 31(1):361–395.

Snijders, T. A. B. (2011). Statistical Models for Social Networks. *Annual Review of Sociology*, 37(1):131–153.

Snijders, T. A. B. and Koskinen, J. (2017). Multilevel Longitudinal Analysis of Social Networks. *Unpublished manuscript.*

Snijders, T. A. B., Lomi, A., and Torló, V. J. (2013). A model for the multiplex dynamics of two-mode and one-mode networks, with an application to employment preference, friendship, and advice. *Social Networks*, 35(2):265–276.

Snyder, T., Dillow, S., and Hoffman, C. (2007). Digest of Education Statistics, 2006. Washington D.C.

Spelke, E. S. (2005). Sex Differences in Intrinsic Aptitude for Mathematics and Science? *American Psychologist*, 60(9):950–958.

Stadtfeld, C., Mascia, D., Pallotti, F., and Lomi, A. (2016). Assimilation and differentiation: A multilevel perspective on organizational and network change. *Social Networks*, 44:363–374.

Stadtfeld, C., Snijders, T. A. B., Steglich, C. E. G., and van Duijn, M. A. (2018). Statistical Power in Longitudinal Network Studies . *Sociological methods & research*.

Steglich, C., Snijders, T. A. B., and Pearson, M. (2010). Dynamic networks and behaviour: separating selection from influence. *Sociological Methodology*, 40:329–393.

Stock, J. H. and Watson, M. W. (2008). Heteroskedasticity-Robust Standard Errors for Fixed Effects Panel Data Regression. *Econometrica*, 76(1):155–174.

van der Vleuten, M., Jaspers, E., Maas, I., and van der Lippe, T. (2016). Boys' and girls' educational choices in secondary education. The role of gender ideology. *Educational Studies*, 42(2):181–200.

van Ewijk, R. and Sleegers, P. (2010). Peer ethnicity and achievement: a meta-analysis into the compositional effect. *School Effectiveness and School Improvement*, 21(3):237–265.

van Rijsewijk, L., Dijkstra, J. K., Pattiselanno, K., Steglich, C., and Veenstra, R. (2016). Who helps whom? Investigating the development of adolescent prosocial relationships. *Developmental Psychology*, 52(6):894–908.

van Tubergen, F. and van de Werfhorst, H. (2007). Postimmigration investments in education: a study of immigrants in the Netherlands. *Demography*, 44(4):883–898.

Van Zalk, M. H. W., Kerr, M., Branje, S. J. T., Stattin, H., and Meeus, W. H. J. (2010). It takes three: Selection, influence, and de-selection processes of depression in adolescent friendship networks. *Developmental Psychology*, 46(4):927–938.

Veenstra, R., Dijkstra, J. K., Steglich, C., and Van Zalk, M. H. W. (2013). Network-Behavior Dynamics. *Journal of Research on Adolescence*, 23(3):399–412.

Vervoort, M. H. M., Scholte, R. H. J., and Overbeek, G. (2008). Bullying and Victimization Among Adolescents: The Role of Ethnicity and Ethnic Composition of School Class. *Journal of Youth and Adolescence*, 39(1):1–11.

Wang, M.-T. and Degol, J. (2013). Motivational pathways to STEM career choices: Using expectancy–value perspective to understand individual and gender differences in STEM fields. *Developmental Review*, 33(4):304–340.

Wang, M.-T., Eccles, J. S., and Kenny, S. (2013). Not Lack of Ability but More Choice: Individual and Gender Differences in Choice of Careers in Science, Technology, Engineering, and Mathematics. *Psychological Science*, 24(5):770–775.

Watt, H. M. G., Shapka, J. D., Morris, Z. A., Durik, A. M., Keating, D. P., and Eccles, J. S. (2012). Gendered motivational processes affecting high school mathematics participation, educational aspirations, and career plans: A comparison of samples from Australia, Canada, and the United States. *Developmental Psychology*, 48(6):1594–1611.

Wentzel, K. R., Barry, C. M., and Caldwell, K. A. (2004). Friendships in Middle School: Influences on Motivation and School Adjustment. *Journal of Educational Psychology*, 96(2):195–203.

Wentzel, K. R., Battle, A., Russell, S. L., and Looney, L. B. (2010). Social supports from teachers and peers as predictors of academic and social motivation. *Contemporary Educational Psychology*, 35(3):193–202.

Wigfield, A. and Eccles, J. S. (2000). Expectancy–Value Theory of Achievement Motivation. *Contemporary Educational Psychology*, 25(1):68–81.

Wölfer, R., Cortina, K. S., and Baumert, J. (2012). Embeddedness and empathy: How the social network shapes adolescents' social understanding. *Journal of Adolescence*, 35(5):1295–1305.

Wölfer, R., Faber, N. S., and Hewstone, M. (2015). Social network analysis in the science of groups: Cross-sectional and longitudinal applications for studying intra- and intergroup behavior. *Group Dynamics: Theory, Research, and Practice*, 19(1):45–61.

Wright, J. C., Giammarino, M., and Parad, H. W. (1986). Social status in small groups: Individual–group similarity and the social "misfit". *Journal of Personality and Social Psychology*, 50(3):523–536.

Youniss, J. and Smollar, J. (1985). *Adolescent relations with mothers, fathers, and friends*. University of Chicago Press, Chicago.

Yuan, Y. and An, W. (2017). Context, network, and adolescent perceived risk. *Social Science Research*, 62:378–393.

List of Figures

List of Tables