

Friendship network formation in Chinese middle schools: Patterns of inequality and homophily

Weihua An ^{*}

Emory University, United States

ARTICLE INFO

Keywords:

Social networks
Social inequality
Status differential
Differential homophily
Adolescent friendship

ABSTRACT

Previous studies on adolescent friendship network formation emphasize the roles played by individual characteristics, structural opportunities, and endogenous tie formation processes. This study extends previous studies in three major aspects: 1) developing and testing theories of status differential (i.e., friendship ties are more likely to run from low status subjects to high status subjects than vice versa) and differential homophily (i.e., homophily is stronger between high status subjects than between low status subjects) through modeling tie-mixing effects of individual characteristics, 2) including qualitative analyses to elaborate selected mechanisms, and 3) comparing selected friendship patterns between China and the U.S. The extended framework is applied to analyzing survey data obtained from 4,094 students in six middle schools in China. The results show that friendship ties are unevenly distributed by individual characteristics (status differential is confirmed for five of the seven characteristics examined), there is significant tie homophily (stronger homophily among high status subjects than among low status subjects is observed for five of the seven characteristics examined with three of them being statistically significant), and endogenous tie formation processes such as reciprocity, transitivity, and preferential attachment are strongly present. Although some of these patterns are common to both China and the U.S., there are also notable differences between the two. Lastly, I discuss the implications and limitations of this study.

Introduction

Social capital embedded in social networks can convey critical instrumental benefits or emotional support (Bian, 1997; Carbonaro and Workman, 2013; Cook, 2014; Dowd and Pinheiro, 2013; Kawachi and Berkman, 2001; Lin, 2001; Pinheiro and Dowd, 2009; Song, 2011; Perry and Pescosolido, 2015). Adolescent friendship is a primary source of social capital in adolescence and an early origin of social capital in adulthood. Adolescents with more friends tend to enjoy stronger social support, have higher social status in their groups, and possess a special socioemotional skill of “being a good friend” that is beneficial to their future careers (Shi and Moody, 2017; Thoits, 2011). Hence, studying adolescent friendships offers a window through which one can observe how socioeconomic inequalities delimit everyday social interactions and how the inequalities are reinforced by social networks (DiMaggio and Garip, 2011, 2012; Small, 2009). Studying adolescent friendships also helps address “the heightened feelings of loneliness” in adolescence (Cavanaugh and Buehler, 2016; Ladd and Ettekal, 2013). Last but not least, studying adolescent friendships is useful for identifying peer

leaders or groups who may be called upon to diffuse positive norms and behaviors (Centola, 2011; Paluck et al., 2016; Valente, 2012).

Previous studies have offered three major perspectives to understand friendship formation. (1) Some focus on the roles of individual characteristics such as age and gender (Gest et al., 2001, 2007; Lewis, 2013). (2) Others focus on structural opportunities such as propinquity and shared activities (Feld, 1981, 1982; Wimmer and Lewis, 2010; Zeng and Xie, 2008). (3) Recent studies incorporate endogenous tie formation processes such as reciprocity and transitivity (Block, 2015; Goodreau et al., 2009; Wimmer and Lewis, 2010). As is shown by Wimmer and Lewis (2010), it is important to account for all three perspectives in order to fully understand and more accurately estimate the multiple competing mechanisms for friendship formation.

The current study makes three contributions to the literature on friendship formation. First, it demonstrates a new analytical approach that models the tie-mixing effects of individual characteristics. When examining the effects of individual characteristics, previous studies mostly focus on whether subjects with a certain characteristics are more likely to receive or send ties and whether ties are more likely to be

^{*} Corresponding author at: Departments of Sociology and Quantitative Theory and Methods, Emory University, 1555 Dickey Dr., GA, 30322, United States.
E-mail address: weihua.an@emory.edu.

formed between subjects with the same characteristics. Previous studies rarely distinguish friendship ties by alter's characteristics. As a result, when studying the receiver effect they cannot tell from whom friendship ties originate and when studying the sender effect they cannot tell to whom friendship ties are directed. They also mingle homophily between subjects possessing a certain characteristics (e.g., smoker-smoker ties) with homophily between subjects without the characteristics (e.g., nonsmoker-nonsmoker ties). To overcome these limitations, in this paper I advocate modeling the tie-mixing effects of individual characteristics. This approach allows one to examine how friendship ties are formed based upon consideration of both egos and alters' characteristics. As we will see later, if a characteristics can be attached with social status value, this approach is also well suited to study whether friendship ties follow status differential (i.e., friendship ties are more likely to run from low status subjects to high status subjects than vice versa), a theory originally proposed by [An and McConnell \(2015\)](#). This approach also allows one to test a theory of differential homophily, namely, whether homophilous ties are more likely to be formed among high status subjects than among low status subjects. Making such a distinction is important because it helps understand how social status interacts with friendship homophily and also helps develop ideas for targeted interventions by status groups. To note, while examining the tie-mixing effects of individual characteristics, this paper also account for the roles of other social processes in tie formation such as structural opportunities and endogeneity tie formation processes.

Second, unlike prior studies that mostly use only quantitative data or only qualitative data, this study takes a mixed-methods approach by combining quantitative and qualitative analyses. Besides providing quantitative analyses of survey data on students' friendships, this study also collects and analyzes students' reflections on their own friend choices and interviews with selected teachers and students. Analyses of the qualitative data help elaborate on selected findings discovered in the quantitative analyses and also reveal some new mechanisms that are not captured by the quantitative analyses. Meanwhile, quantitative analyses of the survey data help expose certain subjectivity and social desirability bias in the qualitative data. Hence, by combining quantitative and qualitative data analyses this study provides more enriched and accurate understandings on friendship formation.

Third, this study offers a cross-society comparative perspective on adolescent friendship formation. Most prior studies are conducted in the U.S. or Europe. However, because the education system and school culture are dramatically different in China, this study offers an opportunity to examine the generalizability of previous findings in a different context. Also, as economic inequality has been rising rapidly in China ([Davis and Wang, 2009](#); [Whyte, 2010a, 2010b](#); [Whyte and Im, 2014](#); [Wu and Treiman, 2007](#); [Xie and Zhou, 2014](#)), whether and how economic inequality in China is manifested in adolescent friendship formation itself is an important question to tackle.

The paper proceeds as follows. I will first review the traditional perspectives on adolescent friendship formation. Then I will elaborate on the contributions of this study in relation to the literature, in particular, how modeling tie-mixing effects helps more accurately test the theories of status differential and different homophily on friendship formation. After that, I will describe the data and methods and present the results. Lastly, I will conclude and discuss the implications and limitations of this study.

Friendship formation: traditional perspectives

There has existed a fair amount of literature on adolescent friendship formation ([Asher and Gottman, 1981](#); [Berndt and Ladd, 1989](#); [Bukowski et al., 1996](#); [Chen et al., 2006](#); [Goodreau et al., 2009](#); [Kupersmidt and Dodge, 2004](#); [Ladd, 2005](#); [Renshaw, 1981](#); [Rubin et al., 2009](#); [Slee and Rigby, 1998](#)). Broadly speaking, past research has offered three major perspectives to understand adolescent friendship formation in school context. [Fig. 1](#) provides a summary of these perspectives and the new ones to be introduced in this paper.

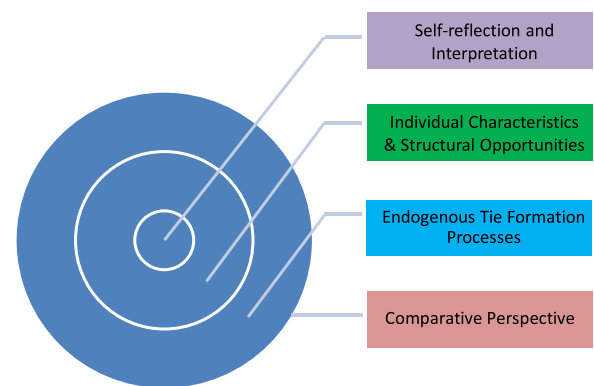


Fig. 1. A Framework for Understanding Friendship Formation.

Individual characteristics

The individual perspective emphasizes the roles played by individual factors such as demographics and socioeconomic factors in friendship formation ([Gest et al., 2001; 2007](#)). Prior research shows that adolescent friendship ties tend to be unequally distributed by these factors. For example, [Gest et al. \(2007\)](#) showed that boys' friend groups were somewhat larger than girls'. [Malacarne \(2017\)](#) showed that in schools with students from diverse socioeconomic backgrounds, higher-SES students received more friendship ties. Another well-established finding in the literature is homophily, namely, friendship ties tend to occur between subjects with similar characteristics ([Block and Grund, 2014](#); [Leszczensky and Pink, 2015, 2017](#); [McPherson et al., 2001](#); [Wimmer and Lewis, 2010](#)). For example, [Leszczensky and Pink \(2015, 2017\)](#) showed that students tended to be friend with classmates, peers with the same ethnicity, as well as peers with the same religion. [Wimmer and Lewis \(2010\)](#) showed that friendship ties tended to occur more between students with the same race, ethnicity, socioeconomic status, musical taste, or regional origin.

When examining the roles of individual characteristics, past studies typically focus on receiver effect (i.e., effect of a certain characteristics on receiving ties), sender effect (i.e., effect of a certain characteristics on sending out ties), and homophily (i.e., the tendency to affiliate with similar others) ([Laumann, 1965](#); [Leszczensky and Pink, 2015, 2017](#); [Wimmer and Lewis, 2010](#)). This modeling approach allows one to examine how a certain characteristics is associated with popularity (i.e., the tendency of receiving ties) and sociality (i.e., the tendency of sending ties). Take smoking as an example. A positive receiver effect for smoking indicates that smokers, as compared to nonsmokers, are more likely to receive ties while a positive sender effect for smoking indicates that smokers, as compared to nonsmokers, are more likely to send out ties. However, this approach does not tell from whom smokers receive their ties. Is it from other smokers or from nonsmokers? Answer to this question helps us better understand how friendship ties are formed based on not just ego's characteristics but also alter's characteristics.

The traditional approach to modeling homophily is also limited. Still take smoking as an example. A positive homophily effect for smoking indicates that ties are more likely to be present between subjects with the same smoking status. However, it does not tell whether homophily is more pronounced among smokers or among nonsmokers. In reality, the two types of homophily can differ greatly and thus it is important to separate them.

In short, previous studies typically do not distinguish friendship ties by alter's characteristics. As a result, they cannot tell exactly from and to whom ties are directed and cannot differentiate homophily between subjects with and without a certain characteristics.

Structural opportunities

Friendship formation can also be affected by structural opportunities

that determine how likely and frequently two people can meet and interact with one another. Broadly speaking, structural opportunities include both propinquity (i.e., spatial closeness) and shared activities or affiliations. Wimmer and Lewis (2010) showed that college students who lived in the same residential halls were significantly more likely to be friends than with those living in different residences. Feld (1981, 1982) argued that people whose activities were organized around the same foci would be more likely to be interpersonally tied. As the chance for people from different backgrounds to meet one another varies, it is important to account for structural opportunities in order to more accurately quantify the effects of individual characteristics on friendship formation. There are two basic approaches to accounting for the impact of structural opportunities in friendship formation. One is to exhaustively control for all the structural opportunities. Obviously, this is difficult to do in any study. As such, prior network research has typically assumed that as long as all the subjects in a network have some chance to meet one or another, then they are provided with an equal opportunity to form friendships and the amount of interaction does not matter. This assumption typically works well for confined settings, like classrooms, workplaces, and small- or medium-sized schools.

Endogenous tie formation processes

People can also form friendships because of endogenous tie formation processes such as reciprocity (i.e., the tendency to reciprocate friendship ties), transitivity (i.e., the tendency to be a friend of a friend), and preferential attachment (i.e., the tendency to be a friend of a popular person). These processes are endogenous because they reflect that people's choices of friends depend on existing friendship ties. Controlling for these endogenous processes is important because it not only helps verify the endogenous mechanisms for friendship formation, but also helps provide more accurate estimates of the effects of other factors. For example, the effects of individual characteristics might become notably smaller once endogenous tie formation processes such as transitivity are controlled for (Wimmer and Lewis, 2010).

To note, the three perspectives outlined above are not necessarily mutually exclusive. For example, some of the structural opportunities such as organizational affiliations and participation in organized activities can be “statistically” treated as individual characteristics. Nonetheless, dividing the factors into three groups helps more clearly conceptualize the determinants of friendship formation. In the literature, there is also a tendency to focus on the first two perspectives while neglecting the third. But as pointed out by Wimmer and Lewis (2010), the latter can be important competing or mediating mechanisms for the former two.

The present study

Below I introduce some new perspectives to help advance our understanding on friendship formation.

Tie-mixing effects and status hierarchy

In this study, I model the tie-mixing effects of individual characteristics to address the limitations in the traditional modeling approach. To

note, tie-mixing effects are more appropriately defined for binary characteristics. Specifically, for a binary characteristics, the reference category can be ties directed from subjects with that characteristics to those without the characteristics. With that reference, there are three tie-mixing effects one can model. (1) The effect for ties to run from subjects without the characteristics to those with the characteristics. (2) The effect for ties to form between subjects without the characteristics. (3) The effect for ties to form between subjects with the characteristics.

As shown in Fig. 2, the tie-mixing effects are well suited for studying how friendship formation is shaped by status hierarchy. The status differential theory (An and McConnell, 2015) views friendship ties as venues for social exchanges that can deliver instrumental or affective benefits. Because high status subjects usually possess more resources (be it wealth or prestige), it is expected that low status subjects are more motivated to seek ties with high status subjects than vice versa. Hence, the status differential theory predicts that friendship ties are more likely to run from low status subjects to high status subjects than vice versa. The status differential theory is rooted in the social exchange theory (Cook et al., 2013; Emerson, 1976; Homans, 1958; Lin, 2001) by viewing social connections as venues for social exchanges or as interaction patterns linked to resource patterns (Lin, 2001: 38). The status differential theory is related to the prestige hypothesis proposed by Laumann (1965), which states that people prefer social ties with high status subjects. The prestige hypothesis corresponds to the so-called receiver effect in traditional network models. Conceptually, however, it is different from status differential, which emphasizes from whom an ego receives ties. For example, the prestige hypothesis predicts that high status subjects will receive more ties than low status subjects while status differential predicts that high status subjects will receive more ties from low status subjects than what low-status subjects will receive from high status subjects. Although both the prestige hypothesis and status differential can be correct, they have different emphases and the latter provides a more concrete understanding on tie directionality. Hence, status differential poses a new concept and requires a new modeling approach for testing. The specification of the first tie-mixing effect shown above helps test the implication of the status differential theory.

The latter two effects in the tie-mixing effects are homophily effects as differentiated by the absence or the presence of a characteristics. In the same spirit of the social exchange theory, if ties to high status subjects are viewed to convey more social capital than ties to low status subjects, one can propose a differential homophily theory that predicts that homophily will be stronger between high status subjects than between low status subjects (which may partly explain the core-periphery structure displayed by many social networks). Previous studies including the like-me hypothesis (Laumann, 1965) or equivalently the homophily hypothesis (Lin, 2001: 39) combine the two homophily effects. In contrast, differential homophily also represents a new concept, pointing out that the degree of homophily can vary by the status value. The specification of the two homophily effects in the tie-mixing effects shown above helps test differential homophily.

In short, modeling tie-mixing effects helps more accurately examine how status hierarchy shapes friendship formation. Of course, not all characteristics is status-charged and sometimes it can be murky to judge

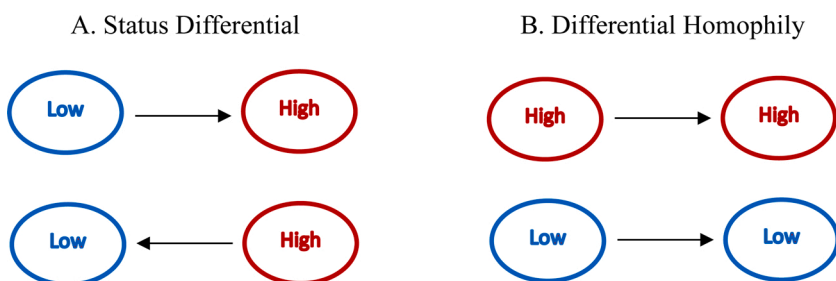


Fig. 2. Tie-Mixing Effects and Status Hierarchy.

Note: Tie-mixing effects are well suited for studying how friendship formation is shaped by status hierarchy. For a binary status characteristics (Low vs. High), a tie can take one of four forms: Low→High, High→Low, Low→Low, and High→High. Comparing the prevalence of the first two forms of ties in a network helps examine whether ties are more likely to follow status differential, i.e., ties are more likely to run from low status subjects to high status subjects than vice versa. Comparing the prevalence of the last two forms of ties helps examine differential homophily, i.e., whether homophily is more pronounced among high status subjects than among low status subjects.

ex ante the status value of a characteristics. Hence, in practice although one should strive to make theory-informed status assumptions, sometimes it may be more appropriate to combine confirmatory analysis for variables with a clear status hierarchy with exploratory analysis for variables without a clear status hierarchy.

Qualitative perspective

Most previous studies on friendship formation rely on quantitative analyses and rarely employ qualitative analyses except a few studies such as Chan Tack and Small (2017). In this paper I combine both quantitative and qualitative analyses. In particular, the qualitative analyses will be based on students' reflections on their own friend choices as well as interview data with selected teachers and students on their interpretations of friendship patterns in their school. We will see that the qualitative analyses help confirm some of the friendship patterns shown in the quantitative analyses and elaborate on the underlying mechanisms. They are also useful for discovering new mechanisms for friendship formation that are overlooked in the quantitative analyses. For example, the qualitative analyses will show why smokers are popular, how teachers and parents can strategically shape students' friendship formation, etc. However, this study does not blindly assume the objectivity of the qualitative data. It will also triangulate them with quantitative data analysis. By doing so, this study can detect some of the subjectivity and social desirability bias in students' self-reflections and interpretations of friendship formation.

Comparative perspective

A number of studies on adolescent friendship formation have been produced in the context of the U.S. (Chan Tack and Small, 2017; Gest et al., 2007; Goodreau et al., 2009; McFarland et al., 2014) and Europe (Block, 2015; Leszczensky and Pink, 2015, 2017; Mercken et al., 2009). But only a handful of relevant studies have been conducted in China (Chen et al., 2018; Lu et al., 2018; Niu et al., 2016). Prior studies, whether conducted in China and other countries, rarely examine how status differential and differential homophily shape friendship formation. Generally speaking, previous studies conducted in China rely on smaller samples which tend to produce less generalizable results. They also include fewer covariates, which can entail larger omitted variable bias on the results. Earlier studies also typically employ conventional regression models while the current study uses exponential random graph models that can capture more nuanced, complex friendship formation patterns. In addition, as mentioned above, this study also provides qualitative analyses to supplement, elaborate, and triangulate with the quantitative analyses. Lastly, many of the earlier studies (Li et al., 2012; Lu et al., 2018) treat friendship as a predictor or a measure of popularity rather than the outcome and thus do not directly analyze friendship formation itself. That said, many of the prior studies are conducted in about the same time period of the current study but in various, different sites in China. Hence, the results may be complementary to the ones shown in the current study.

There are two major reasons for which studying adolescent friendship formation in China is important. First, the topic fills an important gap in research on social inequality in China. Prior studies on social inequality in China (Bian, 2002; Wu and Xie, 2003) mostly focus on outcomes such as education, employment, and income. In comparison, this study views friendship as an important source of social capital and examines how such a source of social capital is unevenly distributed by demographic and social factors. Although studies have examined social causes of inequality in social capital in other populations (Dowd and Pinheiro, 2013; Lin, 2000; Zhang et al., 2021), relevant studies are still rare for adolescents, especially in China. In particular, by using family background to predict student friendship at school, this study shows whether students from more affluent families attract more friendships and whether students from families with different economic conditions

are more likely to cluster in their own circles. As income inequality in China has been growing rapidly (Davis and Wang, 2009; Whyte, 2010a, 2010b; Xie and Zhou, 2014), a high level of inequality in family economic condition is expected to translate into a high level of inequality in adolescent friendships. The result will shed light on one of the mechanisms for intergenerational transmission of social inequality.

Second, studying adolescent friendship formation in China helps provide a comparative perspective on the subject. As most of the relevant research has been conducted in the U.S. and Europe, it is important to know whether and to what extent findings developed in those contexts are applicable to other cultures. On one hand, some of the friendship patterns are expected to be invariant across cultures. For examples, it is expected that friendship ties will generally follow status differential whereby low status subjects are more likely to name high status subjects as friends than vice versa, that homophily in terms of demographics, propinquity, and other social factors will be a main driver for friendship formation, and that friendship ties will follow endogenous formation processes such as reciprocity, transitivity, etc.

On the other hand, differences in status hierarchy and educational institutions can produce differences in friendship formation patterns. For example, racial mixing is a significant research focus in the U.S. whereas it will not be in a largely racially homogenous society like China. Unlike the U.S., students in China mostly interact with their classmates at school and thus friendship ties are expected to be mostly formed within classrooms. In addition, there is substantial opposition for cross-gender interactions in Chinese schools as mandated by social norms (Li et al., 2012). Chinese schools and teachers also emphasize academic performance a lot more (Niu et al., 2016). As a result, students with better grades will have a higher social status and obtain more friendship nominations. In China, smoking tends to serve as a social connection builder (Pan, 2004) or a symbol of maturity (Hesketh et al., 2001). Hence, student smokers are often viewed to have a higher status. China also has a long tradition of masculine culture (Parish and Whyte, 1980) that can generate status premium for male students, although theoretically one should be indifferent about gender difference in social status. In short, this study assumes that boys, older students (for their maturity), taller students (as prior research (Case and Paxson, 2008) shows a strong, positive correlation between height and self-esteem and cognitive ability), student smokers, students with better academic performance, optimistic students, and students from more affluent families enjoy higher social status in Chinese middle schools. These status assumptions are still preliminary, but are instrumental for testing the theories of status differential and differential homophily.

Data and methods

Quantitative data

The data come from the Adolescent Smoking and Network Research (ANSR) that the author conducted with assistance from a local team in 2010–2011 in six middle schools in China. The schools are selected from a site in central China. Although not randomly sampled, based on the school and student characteristics (available upon request), the schools may nonetheless represent a large number of (rural or semi-rural) middle schools in China. The six schools include 90 classes and 4,470 students on the rosters, of which 4,094 (92 %) students filled out a baseline paper-survey. The survey asked students to report their demographics, family background, up to ten closest friends at school, etc. Based on these reports, I constructed a friendship network for each school and a series of student covariates. Because tie-mixing effects are better defined for binary covariates, I dichotomized selected covariates and used them as status indicators. For example, when reporting their family economic condition, students have five options to choose from: 1: Very good, 2: Good, 3: General, 4: Difficult, and 5: Very difficult. I treated the first three responses as “good” while the rest as “not good”. Similarly, for reporting (optimistic) personality, students also have five

options: 1: Very pessimistic, 2: Pessimistic, 3: In-between, 4: Optimistic, and 5: Very Optimistic. I treated the last two responses as “Optimistic” and the rest as “not optimistic”. Specifically, the covariates include gender (1 = boy; 0 = girl), age (1 = older than average classmate; 0 = other), height (1 = taller than average classmate; 0 = other), academic ranking (1 = top 20 in the class; 0 = other), recent smoking status (1 = yes; 0 = no), personality (1 = optimistic; 0 = not optimistic), and family economic condition (1 = good; 0 = not good).

My analyses proceed in two steps. In the first step, I use exponential random graph models to examine the roles of individual characteristics, structural opportunities, and endogenous tie formation processes in friendship formation. In brief, an ERGM assumes that the probability of observing a network is as follows (Goodreau, 2007; Goodreau et al., 2009; Robins et al., 2007; Robins et al., 2009; Snijders et al., 2006; Wasserman and Faust, 1994).

$$\Pr(W = w | X) = \exp\{\theta^T g(w, X)\} / K \quad (1)$$

where W is a random network, w represents the observed network, X the covariates, $g(w, X)$ a list of covariate effects and endogenous tie formation processes, θ the model coefficients, and K a normalizing factor that ensures the probabilities sum to one. More intuitively, an ERGM is equivalent to a conditional logit model (Hunter et al., 2008), namely, the log odds of observing a tie from i to j conditioning on the rest of the network (w_{ij}^c) is determined by $\delta_{ij}(w, X)$, which represents changes in the modeled effects when the i - j tie turns from absence to presence.

$$\text{logit} \left[P(w_{ij} = 1 | w_{ij}^c, X) \right] = \theta^T \delta_{ij}(w, X) \quad (2)$$

I fit two ERGMs for the friendship network in each school. The first model is a conventional ERGM. Its first part includes receiver effects, sender effects, and homophily effects for the covariates. The receiver effect of a covariate, if positive (or negative), reflects students with a particular attribute is more (or less) likely to receive friendship ties. The sender effect, if positive (or negative), reflects students with a particular attribute is more (or less) likely to send out friendship ties. The homophily effect, if positive (or negative), reflects students with the same attribute is more (or less) likely to be friends. The second part of the model includes indicators of whether two students are in the same grade and in the same classroom, respectively, which are used to capture propinquity or structural opportunities for friendships to be formed between students in the same grade or in the same classroom. The last part of the model includes five terms to account for endogenous tie formation processes. The term “edges” (like a regression intercept) accounts for baseline connectivity of the network. The term “mutuality” accounts for reciprocity in friendships. The term “gweswp” (geometrically weighted edge-wise shared partners) accounts for transitivity in friendships, namely, the tendency that a friend of a friend is also a friend. The term “gwdsp” (geometrically weighted dyad-wise shared partners) accounts for the presence of two-paths, namely, the pattern of $A \rightarrow B \rightarrow C$. Because some of the transitive triangles may be created by overlapping two-paths, controlling for two-paths helps more accurately estimate transitivity. I include “gwidegree” (geometrically weighted indegree distribution) and “gwodegree” (geometrically weighted outdegree distribution) to account for heterogeneity in the distributions of incoming ties and outgoing ties. According to Hunter (2007), a negative coefficient for “gwidegree” indicates preferential attachment (i.e., ties are more likely to be directed toward a few subjects) and a negative coefficient for “gwodegree” indicates differential sociability (i.e., ties are more likely to be sent out by a few subjects). Furthermore, Lusher et al. (2013) show that negative coefficients for the two terms indicate that the network tend to exhibit a core-periphery structure.

The second ERGM also includes three parts and is specified specifically for testing theories of status differential and differential homophily. The first part includes tie-mixing effects for the covariates that help test the status differential theory and distinguish differential

homophily. The remaining parts are the same as in the first ERGM (and the results are expected to be the same).

I first used the Monte Carlo Maximum Likelihood Estimation (MCMLE) in “statnet” (Handcock et al., 2003) in R to estimate the ERGMs. As is known, the MCMLE is computationally slow and is hard to converge for big or medium sized networks (An, 2016). The assumption made in the MCMLE that an actor makes a friendship choice with the knowledge of all other friendship ties also seems to be less sensible in the context of these schools. Hence, because of model non-convergence issues and theoretical considerations, in the end I employed a fast approximation method, the Pseudo Maximum Likelihood Estimation (PMLE) (Strauss and Ikeda, 1990; Wasserman and Pattison, 1996), to estimate the ERGMs. The PMLE makes only a local dependence assumption that assumes an actor makes a friendship choice given only the local connections, which is more reasonable in this context. But it has been noted that the PMLE tends to produce larger standard errors for the covariate effects and smaller estimates for the endogenous tie formation processes such as transitivity (van Duijn et al., 2009). Both limitations can lead to more conservative statistical inferences, which is good in the sense that any significant findings shown by the PMLE will likely remain significant if the MCMLE is used for estimation and converges. Hence, I will present the PMLE results in the paper, but additional results based on the MCMLE are available upon request. Selected model diagnostics of the PMLE estimates are shown in Figs. A1 and A2 in the appendix, which indicates a reasonable model fitness.

In the second step, I use the univariate random-effects meta-analysis method (An, 2015; Gasparrini et al., 2012; Snijders and Baerveldt, 2003) to combine the results across schools for each model in order to provide an overarching picture of the results and evaluate variations of the results across schools. The model is shown below, where $\hat{\theta}_{ki}$ is the estimated coefficient for the i th variable in the ERGM for the k th school, θ_i a random intercept with a mean θ and a variance τ_i^2 that is to capture the between-school variation, and ε_{ki} an error term whose variance is assumed to equal to the estimated variance of the i th estimated coefficient in the k th school $\hat{\sigma}_{ki}^2$. The Cochran Q test is used to test whether the estimates are statistically indifferent across schools.

$$\hat{\theta}_{ki} = \theta_i + \varepsilon_{ki}, \quad \theta_i \sim N(\theta, \tau_i^2), \quad \varepsilon_{ki} \sim N(0, \hat{\sigma}_{ki}^2) \quad (3)$$

Qualitative data

In the ANSR baseline survey, students were also asked to list (one or more) reasons for which they have selected their current friends. The reasons students can choose from include: (a) similar physical attributes (e.g., looking, height, weight), (b) personalities match, (c) similar life style and hobbies, (d) friend's family has better economic condition, (e) friend's home is close, (f) friend's desk is close, (g) having common friends, and (h) random factors. The last option is meant to capture non-specific reasons. Among the 4,094 students, 3,857 (94 %) of them provided at least one legitimate response to this question. Although the instrument here is based on the traditional survey method, the information provided is more qualitative in nature because it shows students' reflections on the rationales for their own friendship choices.

The ANSR also employed several research assistants to interview six teachers (including two teachers from each of the three grades) and sixteen students (including six from the 7th and 8th grades, respectively, and four from the 9th grade). Teachers and students were each interviewed for seven to eight questions that ranged from friendship patterns to smoking patterns in their school. For the student interviewees, smokers were disproportionately drawn in order to find more information on student smoking and associated friendship patterns. In total, among the interviewed teachers (denoted as T1-T6), four of them are female and two of them are self-reported smokers. Among the interviewed students (denoted as S1-S16), three of them are female and 13 of them are self-reported smokers.

Table 1
Summary Statistics of Student Characteristics.

Student Characteristics	N	Mean	SD	Min	Max
Incoming ties from schoolmates	4,094	7.24	4.06	0	27
Incoming ties from classmates	4,094	4.76	3.06	0	19
Gender (1 = boy; 0 = girl)	4,094	0.45	0.50	0	1
Age	4,094	13.46	1.14	10	17
Height (cm)	4,094	156.65	8.84	50	200
Ranking (1 = top 20 in the class; 0 = no)	4,094	0.41	0.49	0	1
Smoking (1 = yes; 0 = no)	4,094	0.14	0.34	0	1
Personality (1 = optimistic; 0 = not)	4,094	0.52	0.50	0	1
Family econ cond (1 = good; 0 = not)	4,094	0.82	0.39	0	1

Note: Outliers in the reported age and height are fixed by capping them at reasonable values.

Results

Table 1 presents summary statistics on student characteristics. Note that missing data has been imputed, mostly by treating the missing values as zeros. Some outliers in the reported age and height are fixed by capping them at reasonable values. The smoking rate is about 14 % and almost all smokers are boys. The relatively high prevalence of smoking may be partly due to the generally high prevalence of smoking among Chinese males and partly due to the rurality of the schools.

To help understand the patterns of inequality and homophily in descriptive ways, Table 2 shows the mixing patterns of friendship ties by selected variables. Each cell shows the number of ties from the row characteristics to the column characteristics. The diagonals shows the number of ties that are formed between students with the same characteristics while the off-diagonals the number of ties that are formed between students with different characteristics. The numbers at the diagonals are usually larger than those at the off-diagonals, suggesting there is a high level of tie homophily in almost all characteristics. One exception is smoker-smoker ties, which are less prevalent than other types of ties. For the most part, students without a particular status characteristics are also more likely to send out ties to those possessing that characteristics than vice versa, which indicates that friendship ties tend to follow status differential. Of course, these patterns are based on pure counts of the ties. More rigorous examinations of these patterns will be conducted below by using ERGMs.

ERGM results

Table 3 shows the aggregated results across schools for the first ERGM. The results suggest that boys, taller students, students with higher academic ranking, and smokers are significantly more likely to receive friendship ties than their counterparts at the 1% significance level. For examples, the odds for a tie to be directed to a boy is 1.17 ($=e^{0.16}$) times the odds of the tie being directed to a girl, the odds for a tie to be directed to a higher-ranked student is 1.2 ($=e^{0.18}$) times the odds of the tie being directed to a lower-ranked student, and the odds for a tie to be directed to a smoker is 1.49 ($=e^{0.4}$) times the odds of the tie

being directed to a nonsmoker. Meanwhile, girls and higher-ranked students are significantly more likely to send out friendship ties than their counterparts (both $P < 0.05$). For example, the odds for a tie to be directed from a girl is 1.3 ($=e^{0.26}$) times the odds of the tie being directed from a boy. That boys tend to receive more ties is similar to what is found in the US (Gest et al., 2007; Goodreau et al., 2009; Neal, 2010), but seem to be more pronounced here, which might be a reflection of the masculine culture in China.

There is significant tie homophily in all the covariates except personality. For examples, the odds for a tie to be formed between same-gender students is 3.25 ($=e^{1.18}$) times the odds of the tie being formed between two students with the opposite gender, the odds for a tie to be formed between two students with similar academic performance is 1.23 ($=e^{0.21}$) times the odds of the tie being formed between two students with different performance, the odds for a tie to be formed between two students with the same smoking status is 1.28 ($=e^{0.25}$) times the odds of the tie being formed between a smoker and a nonsmoker, and the odds for a tie to be formed between students from similar family economic conditions is 1.05 ($=e^{0.05}$) times the odds of the tie being formed between students from different family economic conditions. In addition, the odds for a tie to be formed between students in the same grade is about four ($e^{1.32} = 3.74$) times ($P < 0.001$) the odds of the tie being formed between students from different grades and the odds for a tie to be formed between students in the same classroom is about three ($e^{1.07} = 2.92$) times ($P < 0.001$) the odds of the tie being formed between students from different classrooms.

Friendship ties also follow significant endogenous processes. Mutual ties are over eight ($e^{2.16} = 8.67$) times as likely to occur as non-mutual ties ($P < 0.001$). Transitivity in friendship ties is also high and significant. Roughly speaking, the odds for a tie to close a triangle is over three ($e^{1.13} = 3.1$) times the odds for the tie not to close a triangle ($P < 0.001$). The negative and significant estimate on “gwdsp” ($P < 0.001$) indicates the relative absence of two-paths in the friendship network that do not close triangles, which provides additional support for transitivity in friendship. The negative and significant estimates on “gwdegree” and “gwodegree” (both $P < 0.001$) indicate friendship ties are disproportionately more likely to be directed to and from a relatively small number

Table 2
Mixing Patterns of Friendship Ties by Selected Variables.

<i>Gender (Boy = 1)</i>			<i>Age (Older = 1)</i>		
	0	1		0	1
0	15888	1541	0	8458	6444
1	756	11444	1	6573	8154
<i>Height (Taller = 1)</i>			<i>Ranking (High = 1)</i>		
	0	1		0	1
0	8024	6956	0	10702	6059
1	6277	8372	1	5449	7419
<i>Smoking (Yes = 1)</i>			<i>Family Econ (Good = 1)</i>		
	0	1		0	1
0	23165	2802	0	1121	4202
1	1846	1816	1	3975	20331

Note: Each cell shows the number of ties from the row characteristics to the column characteristics.

Table 3
Aggregated ERGM Results on the Friendship Networks (Model 1).

Variables	Est.	SE		Q
Receiver Effects				
Gender (Boy)	0.16	0.04	***	0.06
Age (Older)	−0.03	0.02		0.06
Height	0.07	0.02	***	0.12
Ranking	0.18	0.03	***	0.00
Smoking	0.40	0.05	***	0.00
Personality (Optimistic)	0.01	0.02		0.09
Family Economic Condition (Good)	0.04	0.02		0.86
Sender Effects				
Gender (Boy)	−0.26	0.03	***	0.12
Age (Older)	0.01	0.02		0.17
Height	0.02	0.02		0.69
Ranking	0.04	0.02	*	0.55
Smoking	0.05	0.03		0.37
Personality (Optimistic)	0.02	0.02		0.84
Family Economic Condition (Good)	−0.02	0.02		0.53
Homophily				
Gender	1.18	0.06	***	0.00
Age	0.11	0.02	***	0.60
Height	0.08	0.02	***	0.10
Ranking	0.21	0.02	***	0.14
Smoking	0.25	0.03	***	0.17
Personality	0.02	0.01		0.42
Family Economic Condition	0.05	0.02	*	0.83
Proximity				
Same Grade	1.32	0.13	***	0.00
Same Classroom	1.07	0.04	***	0.00
Endogenous Tie Formation Processes				
Mutuality (Reciprocity)	2.16	0.12	***	0.00
GWESP (Transitivity)	1.13	0.03	***	0.00
GWDSPP (Two-path)	−0.17	0.01	***	0.00
GWIDEGREE (Preferential attachment)	−1.11	0.23	***	0.00
GWODEGREE (Differential sociability)	−1.21	0.28	***	0.00
Edges	−6.32	0.25	***	0.00

Note: The Q column shows the P values of the Cochran Q test for testing effect homogeneity across schools. Significance code: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

of students. The negative and significant estimate on “edges” indicates that the friendship networks are on average sparser than expected by chance alone ($P < 0.001$).

Last, The Q column show the P values testing effect homogeneity across schools. The estimates of all the sender effects are indistinguishable across schools at the 5% significance level. But the estimates of the receiver effects for ranking and smoking, the homophily effect for gender, and all the endogenous tie formation processes differ significantly across schools (all $P < 0.05$).

Table 4 shows the aggregated results across schools for the second ERGM. I focus on the tie mixing effects because other parts of the results are the same as in the first model. The results indicate that friendship ties are significantly more likely to run from girls to boys, from shorter students to taller students, from lower-ranked students to higher-ranked students, from nonsmokers to smokers, and from students from less affluent families to students from more affluent families than vice versa (all $P < 0.05$). If one follows the status hierarchy in Chinese middle schools as outlined in the last section, then there is overwhelming evidence (except for age and personality whose effects are statistically insignificant) supporting the status differential theory.

These results also provide refined patterns about homophily. For example, model 1 shows friendship ties are significantly homophilous in terms of family economic condition. Model 2 shows this homophily can be decomposed into two parts: homophily between students from more affluent families and homophily between students from less affluent families. Although the estimates for the two types of homophily

Table 4
Aggregated ERGM Results on the Friendship Networks (Model 2).

Variables	Est.	SE		Q
Tie Mixing Effects				
Status Differential				
Gender (Girl -> Boy)	0.42	0.06	***	0.15
Age (Young -> Old)	−0.04	0.02		0.58
Height (Short -> Tall)	0.06	0.02	*	0.25
Ranking (Low -> High)	0.14	0.03	***	0.15
Smoking (No -> Yes)	0.33	0.05	***	0.20
Optimistic (No -> Yes)	−0.01	0.03		0.15
Family Econ. Cond. (Not Good -> Good)	0.06	0.03	*	0.40
Differential Homophily				
Gender (Girl -> Girl)	1.43	0.08	***	0.00
Gender (Boy -> Boy)	1.34	0.09	***	0.00
Age (Young -> Young)	0.10	0.03	***	0.20
Age (Old -> Old)	0.08	0.02	***	0.36
Height (Short -> Short)	0.06	0.02	**	0.36
Height (Tall -> Tall)	0.15	0.02	***	0.99
Ranking (Low -> Low)	0.17	0.03	***	0.10
Ranking (High -> High)	0.38	0.04	***	0.01
Smoking (No -> No)	0.18	0.04	***	0.16
Smoking (Yes -> Yes)	0.65	0.08	***	0.01
Optimistic (No -> No)	0.00	0.02		0.82
Optimistic (Yes -> Yes)	0.03	0.03		0.14
Family Econ. Cond. (Not Good -> Not Good)	0.07	0.04		0.81
Family Econ. Cond. (Good -> Good)	0.09	0.02	***	0.99
Proximity				
Same Grade	1.32	0.13	***	0.00
Same Classroom	1.07	0.04	***	0.00
Endogenous Tie Formation Processes				
Mutuality (Reciprocity)	2.16	0.12	***	0.00
GWESP (Transitivity)	1.13	0.03	***	0.00
GWDSPP (Two-path)	−0.17	0.01	***	0.00
GWIDEGREE (Preferential attachment)	−1.11	0.23	***	0.00
GWODEGREE (Differential sociability)	−1.21	0.28	***	0.00
Edges	−6.46	0.23	***	0.00

Note: The Q column shows the P values of the Cochran Q test for testing effect homogeneity across schools. Significance code: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

are not significantly different at the 5% level, the former homophily is found to be stronger and statistically significant by itself ($P < 0.001$) while the latter is statistically insignificant. More analyses (by using one of the homophily effects as reference) show that homophilous ties are significantly more likely to occur between girls than between boys (difference = 0.1, $P < 0.05$), between taller students than between shorter students (difference = 0.09, $P < 0.001$), between higher-ranked students than between lower-ranked students (difference = 0.22, $P < 0.001$), and between smokers than between nonsmokers (difference = 0.45, $P < 0.001$). That girls are more likely to have same-gender ties is contrary to what is expected if assuming girls have a lower status than boys in these schools, but it is consistent with prior findings (Laniado et al., 2016; McPherson et al., 2001; Shrum et al., 1988; Szell and Thurner, 2013). The degree of homophily is statistically indifferent regardless of the value in age, personality, and family economic condition (all $P > 0.05$). The indifference may be partly caused by measurement inaccuracy in these variables (e.g., the way how age is categorized and the muddiness in self-reported and peer-perceived personality and family economic condition). Except for gender and age, overall there is a tendency that homophily is more pronounced between high status students than between low status students. Differential homophily is observed for five of the seven characteristics examined with three of them being statistically significant at the 1% level. Hence, although the support for differential homophily is mixed, it could be more general and robust with better measurement of the variables.

In the ERGMs I specified, it is assumed that students in the same school have some chance to meet each other and potentially to be friends

Table 5

Comparisons of Selected Friendship Patterns between China and the U.S.

	(1) China	(2) U.S.		
Studies	The Current Study	Gest et al. (2007)	Goodreau et al. (2009)	Neal (2010)
Data year	2010	2006	1994–1995	2006
Population	4K students in six middle schools, grades 7–9	427 students in a northeastern school, grades 5 and 7	90 K students in grades 7–12 in nationally representative schools	144 students in a Midwestern urban school, grades 3–8
Methods	ERGM on directed networks	Regression	ERGM on undirected networks	Regression
Friendship Patterns				
Girls	Send out more ties but receive fewer	Smaller network size	Slightly more ties	Smaller network size, but getting bigger at higher grades
Gender homophily	High and slightly higher for female		High	High but slightly lower for female
Other homophily	Age, height, ranking, smoking, family economic condition, grade, and classroom		Grade and race	Race
Differential homophily	Usually higher between high-status actors than between low-status actors			
Other receiver effects	Taller students, higher-ranked students, and smokers			
Status differential	Boys, higher-ranked students, and smokers more likely receive ties from counterparts than vice versa.			
Mutuality	Yes			
Transitivity	Yes	Yes	Yes	
Centralization	Ties are more likely to be sent and received by a small number of students.		Yes	

Note: The U.S. studies presented here are those that have a focus on friendship patterns in middle schools and are selected based on comparability. Because of the covariates and models are different, the comparisons are only suggestive. For U.S. studies on friendship patterns in high schools or other contexts, please see [Chan Tack and Small \(2017\)](#), [McFarland et al. \(2014\)](#), etc.

with each other. The assumption might be too strong as it is possible that some students may not have the chance to meet one another at all. To address the possible violation of this assumption, one solution is to specify all the structural opportunities (e.g., whether two students are in the same classroom, geographic proximity of their homes, as well as shared activities and hobbies) through which two students might be able to meet and then treat the ties between students who have no chance to meet as structural zeros. This approach is infeasible because not all the information is available. Another solution is to provide sensitivity analysis. For example, it may be more reasonable to assume that students in the same grade or the same classroom all have some chance to meet one another. If so, one can fit the ERGMs on the friendship networks by grades or classrooms and then aggregate them, respectively. If the results are similar to those at the school level, then one may conclude that the violation of the assumption does not entail any severe consequence. In other words, this will mean that even if students from different grades or classes are provided with an opportunity to meet, they are still unlikely to become friends. Table A1 in the appendix provides the aggregated results of the ERGMs fitted by grades. As compared to those at the school level, the results are almost the same, both in terms of effect size and significance of the estimates. Hence, the assumption on structural opportunities, even if violated, has little impact on the estimated friendship patterns.

Table 5 summarizes the major friendship patterns found in this study and compares them with those in relevant U.S. studies (selected upon best comparability). Some of the patterns are common to both China and the US. For examples, boys tend to receive more ties, there is strong homophily in terms of gender, and ties can be generated by endogenous tie formation processes such as transitivity and preferential attachment. However, with a richer set of covariates, this paper shows that other characteristics such as height, academic performance, and smoking status are also significant commanders of friendship ties. By specifying tie-mixing effects, this paper is also able to show how friendship ties follow status differential and how tie homophily can be differentiated by the presence or the absence of a status characteristics. To note, these comparisons are provisional because the U.S. studies did not use the same covariates or models as in this paper.

Self-Reflection

Table 6 shows the distribution of the reasons reported by the students for choosing their current friends. First, the results confirm some of the patterns found in the quantitative analyses. For example, over 30 % of the respondents listed “having common friends” as a reason for choosing their friends, which confirms the significant transitivity in friendship ties found in the quantitative analyses. Another 20 % of the respondents listed “similar physical attributes” as a reason, which appears to confirm friendship homophily in terms of gender, age, and height as shown in the quantitative analyses. Second, the results also reveal some new mechanisms for friendship formation that are not captured by the quantitative analyses. For examples, about 70 % of the respondents listed “personalities match” and over 60 % listed “similar life style and hobbies” as one of the reasons for choosing their friends. In addition, 24 % and 14 % of the respondents listed “friend’s home is close” and “friend’s desk is close” as a reason, respectively.

However, the qualitative responses may not be taken for granted blindly. There appears to be some subjectivity or social desirability bias in students’ responses as well. For example, the quantitative analyses show that students are significantly more likely to name students from more affluent families as friends but only about 3% of the students admitted that they chose their friends because of their friend’s family economic condition. In this case, the quantitative results appear to be more credible because the survey question on family economic condition is not directly tied to friend-making and so student are less likely to consciously disguise the link between the two.

Interview results

I group the interview results by the interview questions and by the teachers’ and the students’ perspectives, respectively. Frist is on the major ways through which students met their best friends. According to the teachers, the main ways for students to make friends with each other are through studying together or sharing common personality traits (T4). Also, some students know their friends because they attend elementary schools together (T2). According to the students, there are

Table 6
Self-reported Reasons for Choosing Current Friends.

Reasons	Frequency	Percent
(b) Personalities match	2,611	67.70
(c) Similar life style and hobbies	2,382	61.76
(g) Having common friends	1,324	34.33
(h) Random factors	951	24.66
(e) Friend's home is close	928	24.06
(a) Similar physical attributes (e.g., looking, height, weight)	774	20.07
(f) Friend's desk is close	541	14.03
(d) Friend's family has a better economic condition	100	2.59

Note: This is a multiple choice question, and so the percentages do not add up to 100%.

several ways by which they met their best friends. Some become best friends with their deskmates, i.e., those who share the same desk in a classroom (S4), while others become best friends because they are from the same village (S7). Some find their best friends through extracurricular activities, such as playing basketball and computer games and listening to music (e.g., S2, S3 and S11), while others still keep their best friends whom they met in elementary school (e.g., S10 and S12). For an interesting example, one student recalled, “I met my best friend after a fight with him.” (S16)

Second is on what roles teachers and parents play in shaping students' friendships. Most teachers acknowledge teachers play an important role in shaping students' social interactions and subsequent friendship formation. Five of the six interviewed teachers have taken active measures to facilitate social interactions among their students. Extracurricular activities are one of the strategies the teachers often use. For example, a teacher reported, “There are around 60 students in my class. I divide them into six groups according to their test scores, habits, and special skills. Activities like creating blackboard newspapers and cleaning the classroom are all done by the groups. The leaders of these groups are almost like the vice-heads of my class. Within each of the groups, advanced students will help those who are behind and the optimistic ones will help those who are not so.” (T1) Another teacher (T3) said he had organized tug-of-war competitions, singing competitions, and other social activities and team projects to increase positive interactions among his students. Using model students is another strategy the teachers often rely on to facilitate positive interactions among students. Two teachers (T3 and T4) explicitly pointed out that they had used the students with good academic standing as models for other students to follow.

Some teachers describe that many of their students' parents also take measures to shape their children's social interactions and friendships with peers. For example, one teacher said, “Many parents do not want students to go out on Sundays in order to keep their children from being influenced by other students who have bad behaviors.” (T1) According to the students, however, parents' control over whom they should interact with is often fruitless. For example, one student said, “My parents want me to play with classmates who have good grades. But they do not know exactly whom I play with and whom my friends are.” (S1) Another one recounted, “Parents cannot decide with whom we shall make friends. Neither could teachers specify with whom we play or not to play. I can be friends with anyone I want.” (S9) From the students' perspective, it seems like the students have a high level of autonomy in choosing their friends.

Third is on selection of friends based on smoking. The popularity of smokers seems to be partly driven by a conception that smoking is a symbol of coolness and maturity or a kind of personal skill. One teacher said about the students, “They compare with others to see who is cooler. They believe smoking is cool, and they do not want to change, since they want to build cool images.” (T3) Another teacher echoed this point by saying, “Students compare with each other to see who is more mature. Smoking is marked as a personal skill.” (T4) Similarly, a student said, “Students may compare with other students to see who has the ability to

smoke.” (S11) Smoking-based homophily is also noted. A student said, “Smokers in my classroom often become good friends with each other. Their grades are not good, but they are very active. Their parents also smoke, and they do not supervise their children well.” (S2)

Conclusion and discussion

Previous studies have offered three major perspectives to understand friendship formation that emphasize the roles played by individual characteristics, structural opportunities, and endogenous tie formation processes. In this study, I contribute to this general framework by introducing three new perspectives. First, I advocate a new approach (i.e., tie-mixing effects) to model the relations between social status and friendship formation. For one, this approach helps more accurately examine whether friendship ties follow status differential (i.e., whether low status subjects are more likely to send out ties to high status subjects than vice versa). This approach also helps distinguish homophily between high status subjects and homophily between low status subjects whereas previous studies usually mix the two. The empirical findings in this paper largely support both theories of status differential and differential homophily. However, more work is needed to obtain stronger evidence and to understand the scope conditions.

At the macro level, this study helps provide a comparative perspective on adolescent friendship formation. On one hand, the results of this study confirm some previous findings in studies conducted in the U.S. and Europe, including the roles played by social status, homophily, mutuality, and transitivity in friendship formation. On the other hand, this study also brings our attention to some patterns that might be unique to China. For examples, in China smokers and students with better academic performance tend to enjoy higher social status and receive more friendship ties and students are more significantly likely to form friendship ties with their classmates than others in their school.

This paper also includes qualitative data analyses to supplement the quantitative analyses. Students' reflections on their own friend choices reveal additional mechanisms and factors (e.g., personality match, deskmates, and household distance) that are not captured by the quantitative analyses but are nonetheless important for adolescent friendship formation. Furthermore, the interviews with selected teachers and students help elaborate on the mechanisms for selected findings in the quantitative analyses. For example, smokers are shown in the quantitative analyses to receive more friendship ties but the mechanism is unclear. The interviews suggest that this might be because smoking is viewed by many students as a conduit for social interactions and a symbol of personal capability and coolness. Overall, the combination of quantitative and qualitative analyses helps provide more enriched understandings on the topic.

In terms of theoretical contributions, this study develops theories of status differential and differential homophily. The empirical analyses show that the former receive strong and broad support while the latter receive mixed support. Among the seven characteristics examined, only two, age and personality, that do not follow the status differential hypothesis. Some possible reasons for this include that these characteristics

are not considered as strong social status signals by the students, that the two variables suffer from measurement error, etc. Regarding differential homophily, stronger homophily among high status subjects than low status subjects is observed for five of the seven characteristics examined, including height, academic ranking, smoking, personality, and family economic condition, although only the first three turn out to be statistically significant. Contrary to our expectation, homophily is stronger among girls than boys (statistically significant at the 5% level) and among younger students than older students (however statistically insignificant at the 5% level). This could be due to that girls and younger students tend to possess higher status instead of lower status as it is originally assumed in the study. It is also possible that homophily can be stronger in low status groups than in high status groups if consideration of social solidarity in low status groups outweighs that of social stigma. Overall, the results of this study also imply that socioeconomic inequality can translate into children's network inequality, which may in turn reinforce socioeconomic inequality in the long run.

This study also has several practical implications. One critical challenge in the growing literature on network interventions is to identify key players in a network who can help spread an intervention most widely (Valente, 2012). The typical practice is to select those with the highest network centrality (e.g., those with the most incoming ties) as the intervention targets. However, this study suggests that a more effective approach might be to select key players from different demographic and socioeconomic strata given that there is significant friendship segregation by demographic and socioeconomic factors. Furthermore, the results of this study may be useful for identifying key players even when social network data is unavailable. This study identifies a series of student characteristics that are highly correlated with network centrality. One may use these characteristics (if available) and the estimated coefficients to predict network centrality. Of course, the results are based on data from China, how accurate such predictions are in different contexts should be considered in practice.

This study may be improved in several aspects. First, the measures used in this study are mostly based on self-reports. Although some of these measures, for example, friendship nominations, have conventionally been collected from self-reports, other measures like student academic performance and family economic condition if collected from official records would help strengthen the results. Second, analysis of the cross-sectional data provides correlations, not necessarily causation. There may be other variables that are important but omitted from the statistical models, such as athletic skills, dating behaviors, and aggression. Hence, although the current results seem to be quite robust to omitted variable bias (as partly indicated by the high statistical significance of the estimates), precautions should be taken to not interpret the results in causal terms. Third, only a limited number of interviews are conducted and the student interviewees are disproportionately smokers. Hence, the opinions expressed by the interviewees, especially the student interviewees, may represent only a subgroup of the subjects. Fourth, as demonstrated by research on network ecology (Yuan and An, 2017; Entwisle et al., 2007; McFarland et al., 2014; Small, 2007; Zeng and Xie, 2008), it would be interesting to consider how ecological factors such as school context, teachers' composition, and classroom features can moderate the friendship patterns shown above.

Acknowledgements

The author thanks the Graduate School of Arts and Sciences, the Multidisciplinary Program in Inequality and Social Policy, the Fairbank Center for Chinese Studies, and the Institute for Quantitative Social Science, all at Harvard University, for the support for data collection.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.socnet.2021.07.003>.

References

- An, Weihua, 2015. Multilevel Meta Network Analysis with Application to Studying Network Dynamics of Network Interventions. *Social Networks* 43, 48–56.
- An, Weihua, 2016. Fitting ERGMs on Big Networks. *Social Science Research* 59, 107–119.
- An, Weihua, McConnell, William, 2015. The Origins of Asymmetric Ties in Friendship Networks: From Status Differential to Self-perceived Centrality. *Network Science* 3, 269–292.
- Asher, Steven R., Gottman, John M. (Eds.), 1981. *The Development of Children's Friendships*. Cambridge University Press, New York, NY.
- Berndt, Thomas J., Ladd, Gary W. (Eds.), 1989. *Peer Relationships in Child Development*. Wiley, New York, NY.
- Bian, Yanjie, 1997. Bringing strong ties back in: Indirect ties, network bridges, and job searches in China. *American Sociological Review* 62, 366–385.
- Bian, Yanjie, 2002. Chinese social stratification and social mobility. *Annu. Rev. Sociol.* 28 (1), 91–116.
- Block, Per, 2015. Reciprocity, transitivity, and the mysterious three-cycle. *Soc. Networks* 40, 163–173.
- Block, Per, Grund, Thomas, 2014. Multidimensional homophily in friendship networks. *Netw. Sci.* 2 (2), 189–212.
- Bukowski, William M., Newcomb, Andrew F., Hartup, Willard W. (Eds.), 1996. *The Company They Keep: Friendship in Childhood and Adolescence*. Cambridge University Press.
- Carbonaro, William, Workman, Joseph, 2013. Dropping out of high school: effects of close and distant friendships. *Soc. Sci. Res.* 42, 1254–1268.
- Case, Anne, Paxson, Christina, 2008. Stature and status: height, ability, and labor market outcomes. *J. Polit. Econ.* 116 (3), 499–532.
- Cavanaugh, Alyson M., Buehler, Cheryl, 2016. Adolescent loneliness and social anxiety: the role of multiple sources of support. *J. Soc. Pers. Relat.* 33 (2), 149–170.
- Centola, Damon, 2011. An experimental study of homophily in the adoption of health. *Science* 334, 1269–1272.
- Chan Tack, Anjanette M., Small, Mario L., 2017. Making friends in violent neighborhoods: strategies among elementary school children. *Sociol. Sci.* 4, 224–248.
- Chen, Xinyin, French, Doran C., Schneider, Barry H., 2006. *Peer Relationships in Cultural Context*. Cambridge University Press, New York, NY.
- Chen, Guanghui, Zhao, Qingling, Dishion, Thomas, Deater-Deckard, Kirby, 2018. The association between peer network centrality and aggression is moderated by moral disengagement. *Aggress. Behav.* 44 (6), 571–580.
- Cook, Karen S., 2014. Social capital and inequality: the significance of social connections. In: McLeod, Jane D., Lawler, Edward J., Schwalbe, Michael (Eds.), *Handbook of the Social Psychology of Inequality*. Springer, pp. 207–227.
- Cook, Karen S., Cheshire, Coyce, Rice, Eric R.W., Nakagawa, Sandra, 2013. Social exchange theory. In: DeLamater, John, Ward, Amanda (Eds.), *The Handbook of Social Psychology*. Springer, Dordrecht, pp. 61–88.
- Davis, Debora, Wang, Feng, 2009. *Creating Wealth and Poverty in Postsocialist China*. Stanford University Press.
- DiMaggio, Paul, Garip, Filiz, 2011. How network externalities can exacerbate intergroup inequality. *Am. J. Sociol.* 116 (6), 1887–1933.
- DiMaggio, Paul, Garip, Filiz, 2012. Network Effects and Social Inequality. *Annu. Rev. Sociol.* 38, 93–118.
- Dowd, Timothy J., Pinheiro, Diogo L., 2013. The ties among the notes: the social capital of jazz musicians in three metro areas. *Work Occup.* 40, 431–464.
- Emerson, Richard M., 1976. Social exchange theory. *Annu. Rev. Sociol.* 2, 335–362.
- Entwisle, Barbara, Faust, Katherine, Rindfuss, Ronald R., Kaneda, Toshiko, 2007. Networks and contexts: variation in the structure of social ties. *Am. J. Sociol.* 112 (5), 1495–1533.
- Feld, Scott L., 1981. The focused organization of social ties. *Am. J. Sociol.* 6 (5), 1015–1035.
- Feld, Scott L., 1982. Social structural determinants of similarity among associates. *Am. Sociol. Rev.* 47 (6), 797–801.
- Gasparrini, Antonio, Armstrong, Ben, Kenward, M.G., 2012. Multivariate meta-analysis for non-linear and other multi-parameter associations. *Stat. Med.* 31 (29), 3821–3839.
- Gest, Scott D., Graham-Bermann, Sandra A., Hartup, Willard W., 2001. Peer experience: common and unique features of number of friendships, social network centrality, and sociometric status. *Soc. Dev.* 10 (1), 23–40.
- Gest, Scott D., Davidson, Alice J., Rulison, Kelly L., Moody, James, Welsh, Janet A., 2007. Features of groups and status hierarchies in girls' and boys' early adolescent peer networks. *New Dir. Child Adolesc. Dev.* 118, 43–60.
- Goodreau, Steven M., 2007. Advances in exponential random graph (p*) models applied to a large social network. *Soc. Networks* 29 (2), 231–248.
- Goodreau, Steven M., Kitts, James A., Morris, Martina, 2009. Birds of a feather, or friend of a friend? Using exponential random graph models to investigate adolescent social networks. *Demography* 46 (1), 103–125.
- Handcock, Mark S., Hunter, David R., Butts, Carter T., Goodreau, Steven M., Morris, Martina, 2003. *statnet: Software Tools for the Statistical Modeling of Network Data*. <http://statnetproject.org>.
- Hesketh, Therese, Qu, Jian Ding, Tomkins, Andrew, 2001. Smoking among youths in China. *Am. J. Public Health* 91 (10), 1653–1655.
- Homans, George C., 1958. Social behavior as exchange. *Am. J. Sociol.* 63 (6), 597–606.
- Hunter, David R., 2007. Curved exponential family models for social networks. *Soc. Networks* 29, 216–230.

- Hunter, D.R., Handcock, M.S., Butts, C.T., Goodreau, S.M., Morris, M., 2008. *ergm*: A Package to Fit, Simulate and Diagnose Exponential-Family Models for Networks. *Journal of Statistical Software* 24, 1–29.
- Kawachi, Ichiro, Berkman, Lisa F., 2001. Social ties and mental health. *J. Urban Health* 78, 458–467.
- Kupersmidt, Janis B., Dodge, Kenneth A., 2004. Children's Peer Relations: From Development to Intervention. American Psychological Association, Washington, DC.
- Ladd, Gary W., 2005. *Children's Peer Relations and Social Competence: A Century of Progress*. Yale University Press, New Haven, CT.
- Ladd, Gary W., Ettekal, Idean, 2013. Peer-related loneliness across early to late adolescence: normative trends, intra-individual trajectories, and links with depressive symptoms. *J. Adolesc.* 36 (6), 1269–1282.
- Laniado, David, Volkovich, Yana, Kappler, Karolin, Kaltenbrunner, Andreas, 2016. Gender homophily in online dyadic and triadic relationships. *EPJ Data Sci.* 5, 1–23.
- Laumann, Edward O., 1965. Subjective social distance and urban occupational stratification. *Am. J. Sociol.* 71 (1), 26–36.
- Leszczynski, Lars, Pink, Sebastian, 2017. Intra- and inter-group friendship choices of Christian, Muslim, and Non-religious Youth in Germany. *Eur. Sociol. Rev.* 33 (1), 72–83.
- Leszczynski, Lars, Pink, Sebastian, 2015. Ethnic segregation of friendship networks in school: Testing a rational-choice argument of differences in ethnic homophily between classroom- and grade-level networks. *Soc. Networks* 42, 18–26.
- Lewis, Kevin, 2013. The limits of racial prejudice. *Proc. Natl. Acad. Sci.* 110 (47), 18814–18819.
- Li, Yan, Xie, Hongling, Junqi, Shi, 2012. Chinese and American Children's perceptions of popularity determinants: cultural differences and behavioral correlates. *Int. J. Behav. Dev.* 36 (6), 420–429.
- Lin, Nan, 2000. Inequality in social capital. *Contemp. Sociol.* 29 (6), 785–795.
- Lin, Nan, 2001. *Social Capital: A Theory of Social Structure and Action*. Cambridge University Press.
- Lu, Ting, Jin, Shenghua, Li, Ling, Niu, Li, Chen, Xinyin, French, Doran C., 2018. Longitudinal associations between popularity and aggression in Chinese middle and high school adolescents. *Dev. Psychol.* 54 (12), 2291–2301.
- Lusher, D., Koskinen, J., Robins, G., 2013. *Exponential Random Graph Models for Social Networks: Theory, methods, and Applications*. Cambridge University Press.
- Malacarne, Timothy, 2017. Rich friends, poor friends: inter-socioeconomic status friendships in secondary school. *Socius Sociol. Res. A Dyn. World* 3, 1–13.
- McFarland, Daniel A., Moody, James, Diehl, David, Smith, Jeff, Thomas, Reuben J., 2014. Network ecology and adolescent social structure. *Am. Sociol. Rev.* 79 (6), 1088–1121.
- McPherson, Miller, Smith-Lovin, Lynn, Cook, James M., 2001. Birds of a feather: homophily in social networks. *Annu. Rev. Sociol.* 27 (1), 415–444.
- Mercken, Liesbeth, Snijders, Tom A.B., Steglich, Christian, de Vries, Hein, 2009. Dynamics of adolescent friendship networks and smoking behavior: social network analyses in six European countries. *Soc. Sci. Med.* 69 (10), 1506–1514.
- Neal, Jennifer W., 2010. Hanging out: features of urban children's peer social networks. *J. Soc. Pers. Relat.* 27 (7), 982–1000.
- Niu, Li, Jin, Shenghua, Li, Ling, French, Doran C., 2016. Popularity and social preference in Chinese adolescents: associations with social and behavioral adjustment. *Soc. Dev.* 25 (4), 828–845.
- Paluck, Elizabeth Levy, Shepherd, Hana, Aronow, Peter M., 2016. Changing climates of conflict: a social network experiment in 56 schools. *Proc. Natl. Acad. Sci.* 113 (3), 566–571.
- Pan, Zhenfeng, 2004. Socioeconomic predictors of smoking and smoking frequency in urban China: evidence of smoking as a social function. *Health Promot. Int.* 19 (3), 309–315.
- Parish, William L., Whyte, Martin K., 1980. *Village and Family in Contemporary China*. University of Chicago Press.
- Perry, Brea L., Pescosolido, Bernice A., 2015. Social network activation: the role of health discussion partners in recovery from mental illness. *Soc. Sci. Med.* 125, 116–128.
- Pinheiro, Diogo L., Dowd, Timothy J., 2009. All that Jazz: the success of jazz musicians in three metropolitan areas. *Poetics* 37, 490–506.
- Renshaw, Peter D., 1981. The roots of current peer interaction research: a historical analysis of the 1930s. In: Asher, Steven R., Gottman, John M. (Eds.), *The Development of Children's Friendships*. Cambridge University Press, New York, NY, pp. 1–28.
- Robins, Garry, Snijders, Tom, Wang, Peng, Handcock, Mark, Pattison, Pip, 2007. Recent developments in exponential random graph (p*) models for social networks. *Soc. Networks* 29 (2), 192–215.
- Robins, Garry, Pattison, Pip, Wang, Peng, 2009. Closure, connectivity and degrees: new specifications for exponential random graph (p*) models for directed social networks. *Soc. Networks* 31 (2), 105–117.
- Rubin, Kenneth H., Bukowski, William M., Laursen, Brett, 2009. *Handbook of Peer Interactions, Relationships, and Groups*. Guilford, New York, NY.
- Shi, Ying, Moody, James, 2017. Most likely to succeed: long-run returns to adolescent popularity. *Soc. Curr.* 4 (1), 13–33.
- Shrum, Wesley, Cheek Jr., Neil H., Sandra, Mac D., 1988. *Friendship in School: Gender and Racial Homophily*. Sociol. Educ. 61 (4), 227–239.
- Slee, Philip T., Rigby, Ken, 1998. *Children's Peer Relations*. Routledge, London, UK.
- Small, Mario Luis, 2007. Racial differences in networks: do neighborhood conditions matter? *Soc. Sci. Q.* 88 (2), 320–343.
- Small, Mario Louis, 2009. *Unanticipated gains: Origins of network inequality in everyday life*. Oxford University Press.
- Snijders, T.A.B., Baerveldt, C., 2003. A multilevel network study of the effects of delinquent behavior on friendship evolution. *J. Math. Sociol.* 27, 123–151.
- Snijders, Tom A.B., Pattison, Pip E., Robins, Garry L., Handcock, Mark S., 2006. New specifications for exponential random graph models. *Sociol. Methodol.* 36, 99–153.
- Song, Lijun, 2011. Social capital and psychological distress. *J. Health Soc. Behav.* 52, 478–492.
- Strauss, D., Ikeda, M., 1990. Pseudolikelihood estimation for social networks. *Journal of the American Statistical Association* 85, 204–212.
- Szell, Michael, Thurner, Stefan, 2013. How Women Organize Social Networks Different from Men. *Sci. Rep.* 3 (1), 1–6.
- Thoits, Peggy A., 2011. Mechanisms linking social ties and support to physical and mental health. *J. Health Soc. Behav.* 52 (2), 145–161.
- Valente, Thomas W., 2012. Network interventions. *Science* 337 (6090), 49–53.
- Van Duijn, Marijtte A.J., Gile, Krista J., Handcock, Mark S., 2009. A Framework for the comparison of maximum pseudo-likelihood and maximum likelihood estimation of exponential family random graph models. *Soc. Networks* 31 (1), 52–62.
- Wasserman, Stanley, Faust, Katherine, 1994. *Social Network Analysis: Methods and Applications*. Cambridge University Press.
- Wasserman, Stanley, Pattison, Philippa, 1996. Logit models and logistic regressions for social networks: I. An introduction to Markov graphs and p*. *Psychometrika* 61, 401–425.
- Whyte, Martin, 2010a. Myth of the Social Volcano: Perceptions of Inequality and Distributive Injustice in Contemporary China. Stanford University Press.
- Whyte, Martin, 2010b. One Country, Two Societies: Rural-Urban Inequality in Contemporary China. Harvard University Press, Cambridge MA.
- Whyte, Martin, Im, Dong-Kyun, 2014. Is the social volcano still dormant? Trends in Chinese attitudes toward inequality. *Soc. Sci. Res.* 48, 62–76.
- Wimmer, Andreas, Lewis, Kevin, 2010. Beyond and below racial homophily: ERG models of a friendship network documented on Facebook. *Am. J. Sociol.* 116 (2), 583–642.
- Wu, Xiaogang, Treiman, Donald J., 2007. Inequality and equality under Chinese socialism: the Hukou system and intergenerational occupational mobility. *Am. J. Sociol.* 113 (2), 415–445.
- Wu, Xiaogang, Xie, Yu, 2003. Does the market pay off? Earnings returns to education in Urban China. *Am. Sociol. Rev.* 68 (3), 425–442.
- Xie, Yu, Zhou, Xiang., 2014. Income inequality in today's China. *Proc. Natl. Acad. Sci.* 111 (19), 6928–6933.
- Yuan, Yue, An, Weihua, 2017. Context, Network, and Adolescent Perceived Risk. *Social Science Research* 62, 378–393.
- Zeng, Zhen, Xie, Yu, 2008. A preference-opportunity-choice framework with applications to intergroup friendship. *Am. J. Sociol.* 114 (3), 615–648.
- Zhang, Hongjuan, Han, Rong, Wang, Liang, Runhui, Lin, 2021. Social capital in China: a systematic literature review. *Asian Business & Management* 20, 32–77.