PEER RELATIONS: PEER INFLUENCE, GENETIC SIMILARITY AMONG FRIENDS AND FRIENDSHIP RECIPROCITY

Yilan Fu

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Sociology.

Chapel Hill
2013

Approved by:

Guang Guo

Kathleen M. Harris

Ted Mouw

Francois Nielsen

Yang Yang

©2013
Yilan Fu
ALL RIGHTS RESERVED

ABSTRACT

YILAN FU: Peer Relations: Peer Influence,
Genetic Similarity among Friends and Friendship Reciprocity
(Under the direction of Guang Guo)

The dissertation explores multifaceted nature of peer relations in three chapters: the first chapter uses the natural experiment of randomly assigned college roommates to estimate causal influences of peers on health behaviors. Significant peer effects are found for church attendance, physical exercise, drinking and binge drinking, more importantly, the effects vary by behavioral level of the influencer. The results also show gender differences in peer influence. All findings consistently point to the mechanism that peers influence one's behavior directly by providing opportunities for individuals to engage in the activity. The second chapter extends the line of inquiry on genetic similarity among friends by taking advantages of a quasi-experiment design to test variation in genetic homophily in different social contexts and exploring whether individual choice further gives rise to differentiated genetic similarity through friendship dynamics. Using two independent studies which contain the same set of genetic markers, the study shows that (1) beyond individual genetic polymorphism, friends are more alike than random pairs based on a set of behavior related genes, (2) greater the contextual constrain on friendship choice, smaller the genetic similarity among friends and (3) individual choice may give rise to increased level of genetic homophily. The study suggests the role of genetic similarity in driving the way we pick friends, also highlights the fundamental role of broad social contexts in moderating genetic influences on complex behaviors, such as friendship. In the third chapter, a relatively

understudied area of social networks involves friendship reciprocity, the study investigates the impacts of individual status, reflected by his or her centrality in a social network, on the likelihood of friendship reciprocity between individuals. The study applies both random effect model and discrete choice model to test hypotheses of homophily and status asymmetry. The results consistently support homophily hypotheses. It is found that different from friendship initiation which presents the pattern of both homophily and status asymmetry, friendship reciprocity is majorly driven by homophily. Furthermore, it is suggested that influence domain, popularity, grade and SES might be the major dimensions upon which homophily mechanism functions.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Guang Guo, for his vital guide on my dissertation. I also thank Dr. Kathleen Harris, Dr. Ted Mouw, Dr. Francois Nielsen and Dr. Yang Yang, who provided so many constructive suggestions and kept encouraging me to move the study forward. To Dr. Yong Cai, thank you for all the help and inspiration. I thank my dear Mom and Dad, I always gain strength from your love. And most especially, to my beloved husband S.J., thank you for always putting me as your priority, this could not happen if without your support and love.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

CHAPTER 1

INTRODUCTION

Peer relations had long been the interest of sociologists. The interest is natural given that peers are critical links between individual lives and the larger social contexts, through which the influence from social structure transmits to individuals (D. L. Haynie & Osgood, 2005). Peer relations are important for individual development, and more recently, the scope of study has expanded from dyadic relations to social networks, embeddedness in social networks has increasingly been viewed as determinants of individual and social outcomes. Individual outcomes such as job attainment and advancement (Mouw, 2003); obesity, smoking (Christakis & Fowler, 2007, 2008) and broader social outcomes such as school racial integration (Mouw & Entwisle, 2006) are all found to be greatly impacted by social network positions. Thus, understanding peer relations in both the traditional and the broader senses has become an increasingly important area of inquiry.

The dissertation extends along this line of studies and explores multifaceted nature of peer relations in three chapters. Although, substantive subjects look very different across the chapters: peer influence, genetic similarity among friends and friendship reciprocity, the three subjects all center around the key issue of antecedents and consequences of peer relations and all have a special interest in how social structures moderate these aspects of social relations.

The first two chapters focus on probably the most-studied two aspects of peer relation: peer influence and peer selection. Peer influence looks deceptively simple but technically difficult to measure. The first chapter takes advantage of randomly paired college roommates to study

causal influence of peers on a wide range of social behaviors. The randomized design is arguably the cleanest for studying peer influence as it avoids the question of peer selection which complicates the observational studies(Mouw, 2006). Furthermore, the chapter examines whether behavioral level of the influencer amplifies or attenuates peer influences, and whether peer influence differs by gender.

The second chapter interests in the ubiquitous peer selection phenomenon that people tend to affiliate with others who possess similar socio-demographic backgrounds, behaviors, and attitudes etc. The study asks whether such individual choices have implications for genetic distribution in friendship networks, if so, what are the relative roles of social contexts and individual choices in shaping the genetic distribution?

The third chapter looks at the influences of social network position on individual friendship experiences. Though reciprocity is one of the expectations about affective relations (Laursen, 1993), friendships are not always symmetrical. The unreciprocated relationships always suggest status hierarchies, as Gould (Gould 2002) commented “someone who pays less attention to you than you pay to her implicitly asserts that she is superior to you in status. If you do not respond by withdrawing your attention, you have implicitly agreed. ” The study asks the questions which are stull a relatively open area of inquiry: why some relationships are reciprocated while the others are not? Does it have anything to do with the position one possesses in the peer network?

CHAPTER 2

SOCIAL CONTAGION OF HEALTH BEHAVIORS: EVIDENCE FROM A RANDOMIZED ROOMMATE STUDY

2.1 Introduction

Individuals are interconnected in social spaces. In the past decade, there has been growing awareness of the importance of such interconnectedness, and individual outcomes such as job attainment and advancement (Mouw, 2003), smoking and obesity (Christakis & Fowler, 2007, 2008) etc. are increasingly viewed as determined by the connected social ties. Dyadic relations are the most basic structure of a network and this paper has special interests in the impact of such basic structure on health behaviors.

Previous studies suggest that health behaviors are, in some sense, socially contagious. That is, health behaviors may spread through social ties in a manner like infectious diseases. Social contagion of health behaviors is a kind of person to person contagion of behavioral traits (Christakis & Owler, 2012; Smith & Christakis, 2008). Such findings have great policy implications as it suggests that interventions which mitigate the spread of health risk behaviors across social ties may help address many growing health problems.

While all such findings on social contagion based on observational data suffer from a major bias: peer selection. That is, individuals clearly do not randomly choose their social relationships, especially for those probably most influential relations such as spouse and friends etc. The observed resemblance in behaviors between two people connected by a social tie may be due to peer influence, or the social contagion, while it is equally possible

that similarity in behaviors brings in the two people together to form a social tie in the first place.

This study exploits a randomized design that eliminates the possibility of peer selection. On most US college campuses, in the freshman year, campus housing roommates are randomly assigned based on a few criteria including preferences stated in students' housing application prior entering college (such as campus location and smoker/non-smoker roommate) and gender. This natural experiment design provides a unique methodological advantage for the study to yield unbiased estimates.

Moreover, peers in college provide one of the most promising avenues for studying social contagion of health behaviors with regard to the following aspects: first, entering college is marked by pronounced changes in the context of students' life. Independent from home and parents, college life presents increased autonomy, more unstructured time and diverse social opportunities. College culture also stands out regarding being more approving of health risk behaviors such as binge drinking and substance use (P. B. Johnson, 1989). Individuals entering college show marked increases in alcohol and drug use, compared to those that live at home or get jobs following high school graduation (Johnston, O'Malley, Bachman, & Schulenberg, 2008). Secondly, to adopt the new environment and the new college roles, students, especially in the first year, seek to establish a peer network in which they find a comfortable position and that can be a source of support and intimacy. Peers are the most salient social referents in their college life, they go to class together, hang out in unstructured time and social occasions and socializing with peers leads to even more social opportunities. Thirdly, college year is a transitional stage from adolescence to adulthood, and students learn and negotiate adult roles through interacting with peers. The great importance

of peer group during the transitional stage suggests that college students are particularly susceptible to the impacts of peers.

College roommates represent an important peer setting to study because roommates spend a lot of time together in an environment that requires sharing a living space and involves day-to-day exposure to one another's behaviors. Empirical evidence has supported the potential influence of roommates as well. A 2001 survey of randomly assigned roommates in a college with students from less advantaged backgrounds showed substantial interaction between students and their first year roommates. On average, a student spent 21.66 hours per week with his or her roommate and 21.28 hours per week with the non-roommate that he/she spent the most time with. 47% of students spent more time with their roommate than any other friend. Moreover, although only 37% of students listed their roommate as one of their best four friends, 72% of students spent more time with their roommate than at least one of their three best non-roommate friends (Stinebrickner & Stinebrickner, 2006). Thus, whether they are friends or not, roommates spend a substantial amount of time together out of necessity, it is reasonable to examine freshman roommates if one is interested in looking for evidence of peer effects.

A number of studies, beginning with Sacerdote (Sacerdote, 2001), examine peer effects in the context of randomly assigned college roommates (G. J. Duncan, Boisjoly, Kremer, Levy, & Eccles, 2005; Foster, 2006; Kremer & Levy, 2008; Stinebrickner & Stinebrickner, 2006; Yakusheva, Kapinos, & Weiss, 2011; Zimmerman, 2003). Interests were initially on peer effects on educational outcomes, and more recently extended to health and health behaviors. Although consistent stories linking various effects are absent, all the studies suggest that roommate effects on individual outcomes are likely to be mediated by roommate

behaviors, and stronger and more significant effects from peers should be found by modeling behavioral outcomes (Foster, 2006; Kremer & Levy, 2008; Yakusheva et al., 2011). Furthermore, as literature had not been able to yield strong and consistent evidence regarding health behaviors, this line of inquiry asks for further exploration with larger sample and methodologically robust models.

The present study exams a wide range of health behaviors: self-pray and church attendance[1], physical exercise, drinking, binge drinking, smoking and marijuana use. Taking advantages of the randomized college roommate design, the study answers to the questions: whether health behaviors are contagious among college roommates? Which behaviors are more susceptible to the peer influence and which behaviors are less so? Further, the study asks whether peers influence males and females in different behavioral domains. Moreover, though limited by data availability, the study provides some tentative exploration on persistence of the first year college roommate effects, and discusses possible mechanisms through which health behaviors spread across college roommates[2].

In the following section, the paper provides theoretical backgrounds for peer effects on health behaviors. The paper then discusses methodological issues related to estimating peer effects and proposes an empirical strategy. In the next section, the paper turns to a description of the data used for analyses. Finally, empirical results and discussions are

[1] The link between religion and health has long been the interest not only of the academia but also the general public. Consistent evidence has shown generally desirable effects of religious involvement (e.g. frequency of church attendance) on a wide range of health outcomes including both physical and mental wellbeing. The possible mechanisms for health effects of religion involve regulation of individual lifestyles and health behaviors, provision of social resources, promotion of positive self-perceptions, provision of specific coping resources etc. (Ellison & Levin, 1998). Thus in the study, church attendance is viewed as a health behavior.

[2]This study looks at a wide range of health behaviors and answers to the basic questions: whether the behaviors are contagious among peers? If so, which behaviors are more susceptible to peer influence? Will the influence last long? and whether males and females are influenced differently? Two other studies by Guo (Guo et al., 2012) and Owen (Owen, 2012) used the same dataset while interested in different research questions: Guo only looked at peer influence on drinking and interested in the individual characteristic – prior drinking history as a moderator of the influence. Owen asked whether roommate relationship quality mediates peer influence.

presented.

2.2 Theories

Based on a framework developed by Denise Kandel (Kandel, 1985), peers may influence health behaviors in two ways: directly and indirectly. Direct peer influences focus on offering opportunities for some behaviors such as going to churches, gyms or bars together, or providing cigarettes or marijuana to a friend. Indirect influences, emphasize the way that peers, through their own actions, may provide information about what behaviors are accepted and admired, what is considered appropriate in a given social context, and therefore what behaviors are likely to lead to social acceptance or reinforcement (Borsari & Carey, 2001).

2.2.1 The Direct Peer Influence

Although the direct peer influence mechanism through offering opportunities to engage in certain behaviors is self-explanatory and studies examining this phenomenon are relatively rare, one important theme emerges through reviewing the literature. In college, years in school were found positively correlated with refusing an offered drink (Klein, 1992), the study suggested that maturity and/or social confidence may make students more resilient to peer offers of alcohol. The relationship between school year and acceptance of opportunities offered by peers should be able to be generalized to other health behaviors beyond drinking. New students attempting to develop friendships with peers and adapt to college life may be more likely to accept offers of opportunities from peers. As school year advances, students gained social ease and established group of peers, which provide them more leverages to say no to some offers.

2.2.2 The Indirect Peer Influence

Peers may influence one's behavior indirectly through changing his or her "definition" about the behavior. Various sociological theories, including symbolic interaction theory (Mead, 1934), differential association theory (Sutherland, Edwin Hardin; Cressey, 1955) and social learning theory (Akers, Krohn, Lanza-Kaduce, & Radosevich, 1979), all point to the peer socialization process that the attitudes, values and thus behaviors of individuals are influenced to become similar to whom they associate with.

Differential association theory and social learning theory are prominent in explaining peer socializations especially for crime and deviance. In college settings, excessive drinking, smoking and substance use should be studied under the framework of deviant behaviors. According to Sutherland's (1995) differential association theory, such health risk behaviors are learned through intimate social relations with peers where attitudes or "definitions" favorable to rule breaking. This theoretical aspect emphasizes on the roles of peers in individuals' learning of group norms toward certain behaviors. The social transmission of conventional or deviant behaviors occurs through peer by disseminating attitudes about the appropriateness of certain behaviors (Sutherland, Edwin Hardin; Cressey, 1955). The attitudes toward certain behaviors come from the social context and peers transmit the contextual impacts to individuals. College is the context in which heavy drinking and substance use are approved (Borsari & Carey, 2001), and some students view college as a place to drink excessively in a time limited fashion before assuming the responsibilities of adulthood (Johnston et al., 2008). Individuals learn the "definitions" through socializing with peers which eventually shape their behaviors.

Aker's social learning theory (1979) is an extension to differential association theory

and suggests that peer impact occurs through individual's imitation of peers' behavior or through the observation of its either positive or negative consequences. All these theories agree on the centrality of peer socialization and view peer influence as the way that changes one's attitude and preference toward the behaviors.

2.2.3 The Direct and Indirect Peer Influences

The two theoretical orientations approach peer influence on health behaviors from different aspects. The direct influence perspective focuses on the importance of peers for creating direct opportunities for individuals to engage in certain behaviors. The indirect influence perspective emphasizes that peers matter because of the dissemination of attitudes about the appropriateness of certain behaviors as well as individual's imitation of peers' behavior and observation of its either positive or negative consequences.

The two mechanisms specified by the perspectives may exert influence independently, that is, peers may influence one's behaviors simply by creating opportunities for him or her to engage in the activities without changing his or her preference for the behaviors. Meanwhile, peers may change one's views on the behaviors while without engaging in the activities together. The two mechanisms also may coexist and exert influence together. It is likely that peers may first encourage individuals to engage in the behaviors by offering opportunities, and through participation and interaction, individual's preference about the behaviors changes.

The two mechanisms may also lead to different expectations regarding the domains and persistency of peer influence. If peers matter only through creating opportunities to engage in the behaviors together, the influences are more likely to be observed among behaviors in the social domain, such as going to churches, drinking etc. which are activities

that peers may engage in together. Peers are unlikely to influence the private domain like self-pray. Moreover, the influence may exist only when peers are together, the impacts may disappear when the roommates no longer live together.

If indirect influence matters, either independently or jointly with direct influence, the effects are not limited to the social domains. Regarding the private behaviors such as self-pray, although peers are unlikely to engage in the behaviors together, one's attitude and eventually behavior may still be influenced through discussion with peers and imitation of their behaviors. Also if the attitudes and preference about behaviors are changes, peer influences are likely to last even when roommates depart.

2.2.4 Does Peer Influence Differ by Gender?

Based on above theoretical frame, the study further explores the question: whether peer influence on health behaviors varies across gender. It is well known that males and females associate in different behavioral domains: women emphasized talking, emotional sharing, and discussing personal problems with their same-sex friends, and men showed an emphasis on sharing activities and doing things with their men friends (Aukett, Ritchie, & Mill, 1988), therefore, it is reasonable to speculate that the general mechanisms of peer socialization are similar to males and females, but when it comes to specific behaviors, one gender might be more affected by peers than the other. Actually, gender difference in peer influence might be present for the following reasons.

First, peer influence might be different across gender due to the nature of peer networks which is gendered; exposure to certain behaviors varies by gender. For example, research has long found that exposure to delinquent peers was greater for males than females (R. Johnson, 1979). Other behaviors, such as drinking, smoking and substance use are more

prevalent among males than females. On the other hand, researchers have demonstrated that women participate more frequently than men in religious activities (Simpson, Cloud, Newman, & Fuqua, 2008), therefore, exposure to religious activity is greater for females than males.

Beyond the differential exposure of males and females to some behaviors, social ties may have differential impacts on two genders as well. In general, females have stronger moral values which disapprove deviant behaviors (Mears, Ploeger, & Warr, 1998), as moral disapproval of deviant behaviors varies by gender, females are less susceptible than males to peer influence on these behaviors. By integrating gender specific perspectives, the study intends to develop general but gender specific explanations of peer influence.

2.2.5 Hypotheses

The study intends to test the following hypotheses derived from the direct and indirect perspectives:

Direct Influence

If peers only matter through creating opportunities to engage in the behaviors:

Hypothesis 1.1:

Peer influences are likely to be salient only among health behaviors in social domains such as church attendance, physical exercise, drinking and substance use, but not private behavior such as self-pray.

Hypothesis 1.2:

Peers influences are likely to disappear when peers depart.

Indirect Influence

If the indirect influence matters, either independently or jointly with the direct

influence,

Hypothesis 2.1:

peers may exert influences on behaviors on both social and private domains.

Hypothesis 2.2:

Peer influences are likely to last even when peers depart.

Peer Influence Varies by Gender

Hypothesis 3:

Peer influence on religious activities is more salient among females and peer influence on drinking, smoking, substance use and physical exercise is more salient among males.

2.3 Methods

2.3.1 Data

The study was conducted in a large state university. This is an elite southern university, according to university year book, for fall 2011 enrolled first year class, 90% of the students had a high school GPA greater than 4.0. The analytical sample is consisted of 2094 students who and whose roommates both participated in the study in 2008 spring semester. The study respondents were drawn from enrolled freshmen, sophomores and juniors in the university at the survey time; they compose 30%, 46% and 24% of the sample. These students were freshmen in the fall semester of 2007, 2006 and 2005 respectively.

Most students live in dormitories for the freshmen year. When housing applications were turned in before entering college, the university housing office placed application data into a large database, and then the database was loaded into software program RMS for random matching. Every student who applied for campus housing was randomly assigned a

unique RMS-ID number. After the first student was placed in a room, the RMS program assigned the next student according to RMS-ID order who had compatible gender, smoking status (smoker or not) and type of requested room (gender composition of hall, geographic area of campus) as the first student's roommate. Students who were randomly assigned with roommates have fairly similar background characteristics to the whole student body (Guo et al., 2012). Evidence does show that roommates were all randomly matched up conditional upon gender, smoking status and requested room type: Guo et al (Guo et al., 2012) tested roommate correlations regarding a wide range of background variables and pre-college behaviors and found no more correlations than what would be expected by chance.

2.3.2 Analytical Strategy

Normally, peer influences in behaviors are estimated by regressing own behavior on peer behavior using observational data. However, these estimated coefficients don't always translate into peer influences. As detailed by Manski (Manski, 1993), there are several biases in interpreting coefficients obtained from such estimation.

First, individuals generally self-select into peer groups, which makes it difficult to separate out the selection effect from any actual peer effect. The first source of bias is called "selection bias".

Second, if individuals i and j affect each other simultaneously, then it is difficult to separate out the causal effect that one has on the other. Manski (1993) refers to this problem as the "reflection bias", showing that this reverse causality will result in overestimation of peer effects on one's behaviors.

Third, there might be shared environmental and institutional influences that affect behaviors or behavior changes of both the individual and his or her peers; this is referred as

"common environmental effects", which will lead to overestimation of peer influence as well.

Keeping the possible biases in mind, the current paper studies peer influence in a setting where peers are randomly assigned conditional upon housing preferences. The randomized design provides unique opportunities to deal with the sources of bias in estimating peer influence:

First, random assignment of roommates eliminates the problem of selecting peers with similar behaviors. In observational studies, lagged measures or controlling for observable characteristics are commonly used to deal with selection bias, while these approaches all have considerable limitations. Randomization is arguably the cleanest design for peer influence as it avoid the question of peer selection which complicates the observational studies(Mouw, 2006).

Secondly, the study uses roommate's pre-college behaviors to explain respondent's college behaviors. Using lagged measures of peer's behaviors also allows eliminating reflection bias which is a major problem when using contemporaneous measures. This approach links peer's behaviors prior freshman year with a respondent's behaviors latter in college years and eliminates the biases from reverse causality.

Moreover, using lagged measures of peer's behaviors also helps to deal with biases from exposure to common environmental influences. The pre-college behaviors of the roommate are not impacted by shared environments – such as living very close to the gym -- which potentially lead to correlation in roommate behaviors. Measuring peer's characteristics prior to exposure to the common college environments is arguably a stronger approach than simply controlling for observable environmental variables or using fixed effects to deal with shared environmental influences. Although the study will also control for campus and

dormitory level factors to further eliminate the influence from common environments.

However, there are limitations associated with the lagged measures. First, peer influence occurs in college, it is roommate's college behavior instead of high school behavior that exerts the effect. Though one's high school behavior is highly correlated with his or her college behavior, it is still reasonable to suspect that pre-college behaviors are more planned by schools and parents, while with greater autonomy in college, students can truly reveal their preferences and choices, thus any changes during the transition to college are missed. To address this concern, the model catches the shortest time lag between roommate's behavior and respondent's behavior. The purpose is to reduce the possible behavioral changes by narrowing down the time window between when roommate's behavior was measured and when the peer influence actually took place. College behaviors were measured though retrospectively at two time points (the first fall semester and the past fall semester), the study uses roommate's high school behavior to predict respondent's behavior in the very first semester in college.

Even though, the biases due to lagged measures can be alleviated but not totally eliminated. It is important to acknowledge the possible biases systematically. The direction of bias is associated with pattern of behavioral changes from high school to college. As seen in figure 1, self-pray, church attendance and physical exercise declined as respondents entered into college, while drinking and binge drinking level elevated. For the behaviors that declined in college, lagged high school measures are likely to overestimate peer's behavior in college, and thus underestimate the peer influence. Similarly, for the behaviors that increased in college, lagged high school measures are likely to yield overestimated peer influence.

Thirdly, the analysis of peer influence might be based on a simplified framework in

which own college behaviors depend on roommate prior behaviors, own prior behaviors, own and roommate background characteristics and campus/dormitory environmental influences. Undoubtedly, college behaviors are also influenced by many other factors including parental influences and peers who are not roommates. However, as long as roommate assignment is independent to all these other factors, the simplified framework still can yield unbiased estimates of roommate effects, which is another leverage of the randomized design.

Lastly, the variables used by housing office in assigning roommates (housing preference such as campus location and smoker/nonsmoker roommate, and gender) should be controlled as fixed effects when estimating peer influence. Other observed campus or dormitory level variables should be controlled as fixed effects in the regression model as well. The fixed effects allow estimates to base upon variations within rather than across the "cells" within which randomization took place.

However, we also need to acknowledge that roommate effects are only one component of the total peer influences experienced by a student: students spend many hours per day interacting with other classmates, athletic teammates and friends on campus. The estimates based on roommates alone will be a lower bound on the total peer effects.

2.3.3 Measures

This survey contains a section in which respondents were presented with a list of behaviors and asked the frequency they undertook the activities. The frequency is measure by 6 categories which are never, less than once a month, once or twice a month, about once a week, 2-4 times a week and every day or almost every day. The key outcomes of interest are college behaviors including self-pray, church attendance, physical exercise, drinking, binge drinking, smoking and marijuana use, which were measured based on the frequency during

the first semester in college. For college behaviors, the frequency categories were coded as a continuous measure which counts times on monthly bases.

The explanatory variables of interest are roommate's high school behaviors, which were retrospectively measured based on the frequency during 12 months prior entering college. The same 6 categories were used to measure high school behavior frequency. While roommate' behavior doesn't seem to influence respondents uniformly. The peer impact is contingent upon the behavioral characteristic of the roommates: a very dedicated roommate should have stronger influences on the respondent than a less dedicated one, and the level of dedication is associated with frequency of behaviors: higher the frequency, more dedicated the roommate and greater the potential influence.

Based on the rationale, roommate's high school behaviors were categorized to reflect his or her level of dedication. In order to yield robust results, several categorization approaches were used. First, a 2 level categorization was also applied: above median – if the roommate undertook the activity above the median frequency in high school; and below median – if the roommate undertook the activity below the median frequency. Then, roommate's high school behaviors were re-grouped into 3 levels: frequent – if the roommate undertook the activity at least 2-4 time a week in 12 months before entering college; occasional – if the roommate ever undertook the activity in high school but at a frequency less than 2-4 time a week; and never – if the roommate never undertook the activity in high school.

As the survey was only conducted at one point of time in college, high school information was gathered retrospectively, recall bias might be a concern. High school information from another independent source was used to test reliability of the recalled high

school measures. Prior to college entrance, some students participated in the Cooperative Institutional Research Program's (CIRP's) Entering Student Survey which asked a rich set of student's characteristics including both demographic backgrounds and many behaviors. Fortunately, the behavioral measures in CIRP are very close to our measures in Roommate Study survey which allows the comparison between the behaviors actually reported in High School and those retrospectively reported in college. As only a subsample of our survey respondents participated in CIRP, the CIRP data was only used for quality check instead of major analyses. Using the subsample of respondents who answered to both CIRP and Roommate Study survey, Guang et al (Guo et al., 2012) found that correlations for measures of two sources were all above 0.8 and highly statistically significant, which supported the validity of recalled high school measures in Roommate Study survey.

2.3.4 Regression Model

2.3.4.1 Estimating Peer Influence

Basically, the model regress one's college behavior on roommate's high school behavior, controlled for high school behaviors of one own, age and socioeconomic attributes of both self and roommate. Therefore, conditional upon the respondent's behavior 12 months prior college, the coefficient estimates the effect of roommate's high school behavior on respondent's behavior in college.

Due to the fact that roommates were randomly paired up, the analysis is not biased by unobserved peer selection. Furthermore, as roommate's behaviors were prior to exposure to the respondent and to the same environmental influence, the analysis is also free to reflection and shared environmental effects biases.

But students were allowed to express their housing preferences, similar students may

self-select into dormitories with specific characteristics, for example non-smokers may choose non-smoking dorms, thus, gender, housing preference (smoking status, preference for campus location and gender composition of hall) are controlled as fixed effect to exam peer influence among roommate pairs who expressed identical housing preference, but were randomly paired up by other observed and unobserved attributes.

Lastly, since pairs of roommates enter the analysis twice (once for each respondent), standard regression methods that assume independency of observations can produce biased estimates of the standard errors, and result in increased probability of type-I error. Thus the study adjusts for clustering at the room level by using multilevel mixed model with room level random intercept.

The regression model is as following:

$$Collegebehavior_{ijk} =$$

$$\beta_0 + \beta_1 RoommateHighSchoolBehavior_{ijk} + \beta_2 SelfHighSchoolBehavior_{ijk} +$$

$$\beta_3 controls_{ijk} + u_k + v_{ij} + e_{ijk} \quad (1)$$

Subscripts i,j,k indicate individual, room, and housing preference type. u_k, v_{ij}, and e_{ijk} are unobserved effects at the level of cells based on housing preference, rooms and individuals. [3]

To reduce sample attrition due to missing values, I employed multiple imputation to fill in missing values in background characteristics. This is a conservative approach which was only applied to background controls but not outcome and predictive variables.

[3] The model was replicated by a Tobit regression procedure to handle the substantial number of "zero" responses in the outcome variables, using the GEE model in SAS. The replication only changed coefficients slightly and does not change the significance of findings qualitatively.

2.3.4.2 Estimating Roommate Behavioral Associations

The above model estimates peer influence in a unilateral way with methodological consideration to deal with "reflection bias". While in reality, peers do influence each other in an interactive and reciprocal way. Although subject to biases due to "common environmental influences", behavioral correlations among roommates still provide useful information regarding how similar the roommates are in terms of the health behaviors. It would be especially valuable if the behavioral associations between roommates are estimated overtime, it suggests whether peer influences accumulate or diminish.

Similar to model (1) which estimates causal peer influence, following model is used to estimate roommate behavioral associations.

$$Behavior_{ijk} = \beta_0 + \beta_1 RoommateBehavior_{ijk} + \beta_2 SelfPriorBehavior_{ijk} + \beta_3 RoommatePriorBehavior_{ijk} + \beta_3 controls_{ijk} + u_k + v_{ij} + e_{ijk} \quad (2)$$

Subscripts are the same as in model (1). The coefficient β_1 estimates roommate behavioral associations controlling for prior behaviors of both respondent and roommate.

2.4 Results

Table 1 provides basic demographic characteristics for the analytical sample. Average age of the college students in our sample is 19.42. Females (62.08%) and Whites (65.76%) are overrepresented. 85% of the students are from a family with at least middle class income. The median family income is from $100,000 - $ 150,000, which is 2 to 3 times of the national median household income. 75% of the students have a mother with at least college education. The sample attributes are very close to those for the overall student body (Guo et al., 2012). Male and female subsamples are similar regarding these demographic

characteristics (data not shown).

Table 2 shows descriptive statistics for outcome variables, which were measured as continuous variables based on days per month during the first semester in college, for the overall sample, male and female subsamples. Gender differences in behaviors are salient. In line with prior research, females prayed and attended religious services more frequently and males were more likely to exercise and use alcohol, cigarettes and marijuana.

The study first exams whether having a roommate who had a health behavior of interest in high school increases respondent's level of the behavior in college. As shown in table 3, the study only finds strong evidence for roommate effects on drinking and binge drinking, but no apparent effects for other behaviors. Having a roommate who drank in high school increases college drinking level of respondent by 0.38 day per month. Having a roommate who binge drank in high school elevates binge drinking level of the respondent by 0.37 day per month respectively.

For most high school behaviors, the Yes or No categories based on whether ever had the behavior in high school contain fair amount of cases. But for physical exercise, a very small proportion, only 2.35%, of the students (49 cases) answered never exercised in high school. Given the small number of cases in the reference category, the study applied another categorization for independent variable with the purpose to check robustness of the above findings. High school behaviors then were classified as in two categories "above median" – the half sample whose behavioral level is above the median level and "below median" – the other half sample whose behavioral level is below the median level. For drinking, binge drinking, smoking and marijuana use, as approximately half or over half sample answered never had the behavior in high school, thus for the two variables the categorization based on

median is the same as yes-or-no categorization.

There are small differences in results using the two categorizations (please see table 4). However, peer influence on physical exercise was revealed after using the below or above median categorization. Respondent's college physical exercise level increased 0.78 days per month by having a roommate who exercised at the above median level in high school. It suggests that the yes-or-no categorization may conceal effects that are contingent upon behavioral level of the influencer.

Though being understudied, studies had shown that characteristics of the influencing peer may magnify or mitigate peer influences (Brechwald & Prinstein, 2011). It is reasonable to predict that peer influence varies according to the level of influencer's behavior. A very dedicated actor should have a greater impact on people around than a less dedicated one. Thus the study re-classified the explanatory variables of interest into three categories in order to capture the behavioral characteristic of the influencer: frequent – if the roommate undertook the activity at least 2-4 time a week in 12 months before entering college; occasional – if the roommate ever undertook the activity in high school but at a frequency less than 2-4 time a week; and never – if the roommate never undertook the activity in high school. The 3-group categorization revealed the peer impacts on religious activities, as shown in table 5, having a roommate who went to church frequently in high school would increase one's college church attendance by 0.42 days per week. While no peer effect was observed if the roommate only went to church occasionally. Similar result was found about physical exercise. Peer influence was only found when the influencer exercised frequently in high school, but not for the influencer who only exercised occasionally in high school. Respondent's college exercise level increased 0.9 day per week when paired with a frequent

high school exerciser. But for drinking and binge drinking, peer influence was only salient when paired with a roommate who drank or binge drank occasionally but not frequently in high school. This finding resonates with the well-known effects of college drinking that moderate drinking is socially desirable while excessive drinking could be disruptive to peers.

The study then exams peer effects for males and females respectively. Roommate's drinking is a powerful predictor of respondent's drinking. Both males and females paired with a roommate who drank occasionally in high school drink more frequently in college than respondents paired with a roommate who didn't drink in high school. Peer effect in physical exercise is only present among males, and peer effect in church attendance is only present among females. The finding is in line with our expectations. Male and female are affected by peers in different behaviors because of their differential exposure to the behavior and the differential influences of gender specific social ties. It is known that females attend religious activities more frequently than males and males exercise more than females. Female same sex friends emphasize on talking, emotional and spiritual sharing and men are more interested in doing activities together with same-sex friends. Moreover, it may also have to do with the type of physical exercise in which male and female participate, males are more likely to participate in games which are group activities, while females may do individual activity more frequently such as jogging.

The study then asks the question whether peer effect persists over time[4]. If peer influence lasts, given the reciprocal nature of peer socialization, we would expect to see increased behavioral correlations among roommates over time. While the results from table 8

[4] To test whether peer influence lasts over time, one way is to set outcome variables to a later time point in college. I tested the model using roommate's high school behaviors as predictors of respondent's behaviors in the semester prior the survey, but barely found any significant results. This suggests that peer influence may not last over the freshman year, while the test may also be underpowered for estimating the smaller sample who were at least in sophomore year (n=1454).

suggest the opposite: for roommates who still lived together after the first year, their behavioral correlations decreased overtime; for roommates who departed after the first year, their behaviors were no longer correlated when they departed. Two themes emerge regarding the findings: first, the behavioral association is strongest in the first semester which is the period of time when freshmen try to adapt to college life and build up friendship networks, and thus are more susceptible to peer impacts. Second, roommate influence was situational, behaviors of roommates synchronized only when they lived together, and their behaviors departed when they no longer lived together[5].

2.5 Conclusion

The study of peer influence using a randomized college roommate design suggests that health behaviors are contagious among college students, more specifically, the analysis yields findings with 4 themes: first, not all behaviors are uniformly influenced by peers, peer effects are significant for church attendance, physical exercise, drinking and binge drinking, but not for self-pray, smoking and marijuana use. Second, peer influence is contingent upon behavioral characteristics of the influencer. For church attendance and physical exercise, the peer influence is only salient when the influencer engaged in the activity very frequently. For drinking and binge drinking, moderate drinker may elevate drinking level of the roommate but the effect disappears for excessive drinker. Third, males and females are influenced by peers in different behavioral domains, though both gender are likely to be influenced by peers in drinking, females are more susceptible to peer impacts in spiritual domain such as church attendance and males are more susceptible to influence in activities such as physical exercise. Forth, the study provides tentative support that peer influence doesn't seem last when

[5] Regardless whether roommates departed after the first year, their behaviors were correlated in the first year, thus the possibility of reverse causality – roommates dissimilar in the first place were more likely to depart after the first year, and thus their behaviors were no longer correlated when they no longer liver together– should be ruled out.

roommates depart, and behavioral correlation among roommates is the strongest for the first semester when they live together.

Theories suggest two mechanisms through which peers influence one's behavior. The direct mechanism suggests that peers influence one's behavior directly by providing opportunities for individuals to engage in the activity. The indirect mechanism suggests that by peers may disseminate attitudes about the appropriateness of certain behaviors, and individuals may change their preferences and behaviors accordingly.

All the results seem more consistent with the direct mechanism. Several observations point to the direction. First, the peer effect is concentrated in social behaviors which may engage other people as companions. For individual behavior such as self-pray, peer impact is not observed. Under the indirect influence hypothesis, people's attitude may be influenced by peers regardless whether they engage in the activity together or not, thus both social and individual behaviors are equally susceptible to peer influence, but this does not appear to hold true. Second, the behavioral association among roommates does not persist when they no longer live together, which is consistent with the direct influence mechanism that roommate changes one's behavior by offering opportunities to engage in the activity together, when the opportunity disappears, the peer influence doesn't sustain any longer.

Peer influence is contingent upon behavioral characteristics of the influencer. By introducing the new dimension, the study revealed peer influences in religious activity and physical exercise which were highly concentrated among influencers who were frequent actors.

The peer effects for drinking and binge drinking and null findings for smoking and marijuana use are consistent with prior studies (G. J. Duncan et al., 2005). The significant

results for drinking and binge drinking are not surprising given the highly social nature of the behaviors among young people, especially college students. For smoking and marijuana use, although large peer effects were estimated for secondary school students in other studies, we didn't find such effects among college students. The discrepancy between our study and prior research could be due to several reasons: first, there could be differences in true effects between secondary and college settings. Study has shown that college selectivity reduces tobacco and marijuana use but has small and possibly positive effects on binge drinking (Fletcher & Frisvold, 2011), that's why college peers may have diminished influences in smoking and marijuana use but magnified influences in drinking or binge drinking. Second, the absence of peer influence in smoking and marijuana use may be due to underreport of the socially undesirable behaviors and small statistically power given the limited number of respondents reported the behaviors. Third, there might be overestimations of peer influences in the studies in secondary school setting, where natural experiments such as random roommate assignments are not available.

The analysis of role of peer influences on a wide range of may have practical implications for university administrators and more general public health efforts. Although the study finds variations in peer influences regarding types of behaviors, behavioral level of the influencer, gender and time period, we tend to interpret the results with caution due to generalizability of one case. The study may bring up more interesting questions than it answers. Future studies with larger sample and in other randomized settings would be worthwhile to see if the results are robust and generalizable across contexts.

CHAPTER 3

SOCIAL CONTEXT AND INDIVIDUAL CHOICE: GENETIC HOMOPHILY AMONG FRIENDS

3.1 Introduction

Homophily, introduced by Lazarsfeld and Merton in 1954 to describe the similarity in attitudes, attributes, and behaviors among close friends, is one of the strongest patterns of human association and is important to investigate because, in part, it is linked to reproduction of social attributes such as concentrated socioeconomic advantages or disadvantages and restricts the flow of ideas, norms and other resources across society (McPherson, Smith-Lovin, & Cook, 2001)

People's personal networks are homogeneous with regard to many sociodemographic (socioeconomic status, race, education and age), behavioral (drinking, smoking, delinquency, depression) and intrapersonal (sociality, popularity, introversion, extraversion) characteristics (Kandel, 1978; McPherson, Smith-Lovin, & Cook, 2001; Kossinets & Watts, 2009; Schaefer, Kornienko, & Fox, 2011; Rivera, Soderstrom, & Uzzi, 2010; Goodreau, Kitts, & Morris, 2009). Evidence also suggests that homophily at genetic level could be observed in friendship networks (Boardman, Domingue, & Fletcher, 2012; Daniels & Plomin, 1985; Fowler, Settle, & Christakis, 2010; Guo, 2006) .

Genetic similarity among friends has long been the interest of social scientists and behavioral geneticists. The inquiries were driven by the belief that though very culturally contingent many social behaviors are, they potentially have a "genetic component"

(Turkheimer, 2000). Similar to many other social behaviors, the search for genetic components of friend making started in the hybrid discipline behavior genetics. Based on pairs of various level of known relatedness, such as identical twins, fraternal twins, full siblings, adoptive siblings …., even without direct measures of genes, behavioral genetics methodology allows to estimate the extent to which variation in a given trait is due to genetic, as opposed to environmental, differences. Compared adoptive and non-adoptive siblings, Daniels and Plomin (Daniels & Plomin, 1985) found that biological siblings were more similar to each other regarding types of friends than adoptive siblings. Similarly, Guo (Guo, 2006) looked at friends of monozygotic and dizygotic twins and found a greater resemblance between friends of monozygotic twins than that between friends of dizygotic twins and full siblings, regarding traits such as GPA, Peabody Picture Vocabulary Test, aggressive behaviors and depression.

Recent advances in molecular genetics made unraveling of the genetic black boxes in traditional behavioral genetics possible. Candidate gene association (CGA) studies, the study on statistical associations between specific genetic variants and behaviors, have gained increased popularity. The trend was facilitated by the increasing availability of both large data sets that contain behavioral data and genetic data and statistical methods. A noteworthy CGA study is from Fowler et al (Fowler, Settle, & Christakis, 2010). It was the first to explore molecular genetic correlations among friends, and they found a positive correlation at DRD2 gene, which has been previously associated with alcoholism, and a negative correlation at CYP2A6 gene, which was associated, though controversially, with openness, among adolescent friends. This study gained the most media interests and also probably the most academic criticisms among all genetic inquires on friendship in social science. While

this line of inquiry is important, as the authors argue, “homophily and heterophily in friendships, expressed at the genetic level, may have notable implications for our understanding both of the way that our genes can shape our environmental exposures and the way that our social environment can shape our behavior.” (Fowler et al., 2010)

Based on the study of Fowler et al, Boardman et al (Boardman et al., 2012)asked a further question: how does social context shape genetically similar friendships? Using the same dataset, the study replicated Fowler et al’s DRD2 gene result and further found that genetic association for DRD2 gene was stronger in schools with greater levels of inequality. The authors argued that this findings suggest that “individuals with similar genotypes may not actively select into friendships; rather, they may be placed into these contexts by institutional mechanisms outside of their control”.

The two studies are of great importance, as for the first time, the “black box” regarding the specific genes based on which friends are connected was unraveled. Fowler et al provided a baseline regarding the positively and negatively correlated genes among friends, Boardman et al brought up a more sociological question: if friends are disproportionately likely to share certain similarities in their genes, does it suggest that genetic similarity drives the way we pick our friends, or other social factors drive us to become friends with people we happen to share genes with?

However, both of the studies have several methodological caveats, which asks for further exploration of the above questions. For Fowler et al’s study, the genetic markers in replication dataset which were used to validate findings from discovery data were imputed instead of directly measured; the negative correlation at CYP2A6 gene was likely due to extremely low variation at this gene (<5%) which resulted in tenuous statistical estimates; the

approach to control for population stratification is less effective and obsolescent; and the methods used to construct friend sample which equate friendship to friend nominations raised concerns as well. However, the most criticized was their approach to look at only six genes, and two out of the six genes were found significant seemed too good to be true. Such criticisms were all raised with the complexity of gene-behavior relationships in concern. Most human traits with a genetic component are **polygenic** (Lewis, 2003), that is, the traits are influenced by a vast number of genes with small effect instead of a few genes with large effects, and when it comes to specific genetic variants, the effects should be far less apparent. As Plomin and Davis (2009) noted, "GWA studies suggest that for most complex traits and common disorders genetic effects are much smaller than previously considered … This finding implies that hundreds of genes are responsible for the heritability of behavioral problems … and that it will be difficult to identify these genes of small effect".

At the same time, the same gene may simultaneously influence many different behavioral traits, which is called **pleiotropic**. The concept endophenotype was used to explain the multifaceted influence of genes. An endophenotype is an intermediate trait or an internal phenotype that lies between genotype and phenotype (Gottesman and Shields 1973). Endophenotype may give rise to an array of phenotypes that all share something with the more primary endophenotype. Thus, given the polygenic and pleiotropic nature of genetic influences, when it comes to a specific polymorphism, its association with friendship should be far less apparent. A more sound approach to explore the baseline for genetic correlations among friends should be looking at a set of genes, which simultaneously influence a wide range of behavioral traits that set up stage for friendship process.

For the work of Boardman et al, as it is a replication and extension of Fowler et al's

study, it shares many issues of prior study, such as only looking at the DRD2 variant, the obsolescent approach for population stratification, the way friend pairs were constructed etc. While a greater concern is about their school environment measures which were hypothesized to moderate genetic correlation among friends. The paper examined two school level characteristics, inequality, measured as Gini coefficient of mother's education) and racial segregation of friendship. Results showed that the racial segregation measure was unrelated to genetic correlation among friends, while the inequality measure moderated the correlation: genetic homophily was not observed in schools with low level of inequality, the strongest genetic homophily emerged in schools with highest level of inequality. First, all the findings relied on only one interactive effect, furthermore, the authors stated "the substantive interpretation of this interaction is that schools with greater levels inequality seem to be schools where students are more likely to make friends with those of a similar DRD2 genotype. One causal mechanism that could lead to such a finding would be that highly unequal schools tend to institute academic tracking policies. Grouping by ability which may induce grouping by race and/or genotype, will constrain the type of 'potential friends' because one is more likely to be friends with classmates than with students with whom one does not share time in the same class". The mechanism is speculative rather than tested, even the mechanism is true, inequality is a poor proxy for school tracking, and thus school context variables of the study, Gini coefficient and racial segregation index, are very limited in measuring the true mechanisms that bring friends together .

The current study shares and expands the interests of Fowler et al and Boardman et al, and asks the following two questions: with the expanded understanding of gene-behavior relationships in mind, the current study focuses on a set of genetic polymorphisms, which

have been found in association with many human behavioral traits, and investigates their overall correlations among friends. The study first answers to the question, based on the set of behavioral related genes, whether genetic correlation among friends exceeds the level expected by chance. More interestingly, the study explores the mechanisms that account for genetic homophily. Does individual choice drive the way people pick genetically similar friends? Or social factors shape the probability of friend genetic homophily?

Two independent studies with the same set of genetic markers while very different social settings for friendship process provide unique leverages for the study to explore above questions. The National Longitudinal Study of Adolescent Health (Add Health) and the Roommate Study contain the same set of 100 genetic polymorphisms from 21 behavior related genes. But their social contexts which shape friendship process are largely different: for Add Health, adolescents were free to name their school friends; for Roommate Study, college roommates were randomly paired up in the first place and they may become friends or not latter when they lived together. The two studies set up the contextual contrast regarding opportunity constrains placed upon students while choosing friends.

The current study contributes to this line of inquiry from the following aspects: first, it expands the genetic scope by looking at a set of genetic polymorphisms simultaneously; second, the different settings of AddHealth and Roommate Study provide unique opportunities to test how social contexts moderate genetic homophily; third, the study distinguishes reciprocal and unilateral friends which provides better measure of friendship and further answers to the unstudied question whether friendship dynamics, such as individual choice on friendship reciprocation, give rise to differentiated levels of genetic similarities among different types of friends; forth, the study improves methodology of prior

studies and uses up to date and more efficient approach to control for population stratification.

3.2 Why Genetic Homophily But Not Heterophily

The studies on homophily can be traced back as early as 1950s, Lazersfeld and Merton's (Lazersfeld & Merton, 1954) analyzed friendship selection among adults and used the term "homophily" to refer to "a tendency for friendships to form between those who are alike in some designated respect" . This landmark work gave birth to a rich tradition of research that investigated the tendency toward similarity in a variety of social settings. Meanwhile, it is well acknowledged that social connections may be between individuals who are simultaneously similar and different; few people do not differ in at least some dimensions and match in at least a few others. As pointed out by Blau (Blau, 1974) in the early 1970s, either homophily or heterophily, the tendency for friendship to form between people who are different in some designated respect, might be an oversimplification for social relations, in some instance, attachment is encouraged by homogeneity across numerous dimensions (McPherson & Ranger-Moore, 1991); in other instances, individuals may seek a balance between similarity on some dimension and difference on others.

Given the multifaceted nature of human relations, why only look at genetic homphily? The study would argue that intimate relation like friendship is exact the type of relationship whose attachment is encouraged by homogeneity, instead of heterogeneity, across numerous dimension, and the benefits of homogeneity increases with the strength of the relationship (Blackwell & Lichter, 2004; Burgess & Wallin, 1943; Qian & Lichter, 2007).

Compared with intimate relations, neutral ties, especially task oriented ties may be more prone to heterophily. Casciaro & Lobo (Casciaro & Lobo, 2008) found that a key fact of organizational life is that people seek out others who they believe to have valuable and

complementary task-related skills. Even in contexts where diversity is explicitly valued and encouraged, homophily still shapes attachment. Ruef et al (Ruef, Aldrich, & Carter, 2003) found that the composition of new business start-ups is driven by similarity in gender, ethnicity, and occupation rather than by functional diversification. Casciaro & Lobo's (Casciaro & Lobo, 2008) field research also suggests that people preferentially collaborate with others who have complementary specializations but similar demographic traits that facilitate communication and trust.

Furthermore, heterogeneous relationships might be instrumental in nature. Collaborative ties between individuals with complementary attributes are often short term and oriented toward the completion of goal. Once the task is completed, the relationship is ended and successful collaborators depart gracefully (Rivera, Soderstrom, & Uzzi, 2010). All above explains why homophily is so ubiquitous in social networks even given the value and necessity of heterophily.

3.3 Mechanisms of Genetic Homophily

Intuitively, the answer for why friends are genetically alike may seem obvious: as illustrated in figure 2, resemblance at phenotypic level, such as behavioral traits and bio-ancestry related traits, may induce friendship. If these observed phenotypic traits have genetic bases, friends should resemble each other at genetic level as well. While resemblance on phenotypic traits may not always equate to genetic correlations, it is well known that homophily on many behavioral attributes may occur through both friendship decision and peer influence (Kandel, 1978), if such phenotypic similarities were solely caused by peer influence, genetic resemblance would not emerge. Thus genetic correlation among friends rises and only rises through friendship decision.

Regarding either behavior or bio-ancestry related genes, friendship decision may give rise to genetic homophily, however the mechanisms leading to homophily might be different. For genes related to behaviors such as drinking, smoking, aggressiveness etc., they may play active roles in friendship selection with the purpose to maximize gene expression, while for bio-ancestry related genes, the homophily is just the unintentional consequences of friendship decision based on race, which is related to physical traits and has social structural implications.

As the study interests in possible roles of genes in friendship decision, the following part will focus on behavior related genes. Two theoretically distinct mechanisms – choice homophily and induced homophily (McPherson and Smith-Lovin 1987), corresponding to the active and passive views of genes – explain the rise of such genetic homophily in social networks.

3.3.1 Choice Homophily

If people prefer to form and retain friendship ties with similar others, genetic homophily could be a result of individual choice, which is called choice homophily. Why people prefer to interact with similar others? This question can be answered from many aspects: psychology emphasizes on ease of communication, shared cultural tastes and other features between similar individuals that smooth the coordination of activity and communication (McPherson, Smith-Lovin, & Cook, 2001). From economic point of view, the cost and risk associated with forming new ties are expected to be lower for similar than dissimilar counterparts due to the shared attributes and experience, correspondingly, the ongoing cost of maintaining ties is also expected to be lower between similar than between dissimilar friends (Kossinets & Watts, 2009). All these aspects hit on one key point that

similar others are expected to provide more benefits than dissimilar ones, which is exactly the genetic underpin for individual preference for similar others according to evolutionary geneticists.

3.3.1.1 Inclusive Fitness and Genetic Similarity Theory

From the genetic view, evolutionary success ultimately depends on passing down maximum copies of one's genes in the population. Genes could be passed down through one's own offspring, which is individual fitness. But individual fitness can't explain altruistic behaviors. Altruism as behaviors carried out to benefit others, which in extreme form involves self-sacrifice, may diminish individual fitness because it runs with the risk of reduced success in passing genes down to offspring. If individual fitness is the only mechanism that works in natural selection, altruistic individuals would eventually be selected against and die out in the population. The concept inclusive fitness was first brought up by Hamilton in 1964 (Hamilton, 1964) to explain altruism: as close relatives share some identical genes, a gene can also increase its evolutionary success by promoting the reproduction and survival of these related or otherwise similar individuals, which is inclusive fitness. Thus individuals maximize their inclusive fitness rather than only their individual fitness by increasing the success of passing down shared genes. Inclusive fitness theory suggests the benefits from others who share genes with you for the evolutionary success of your own genes. Thus, individuals may favor genetically similar others and provide altruism. Study has shown that the self-reported likelihood of risking one's life to save other's life is a function of the degree of genetic relatedness between the help provider and beneficiary (Burnstein, Crandall, & Kitayama, 1994). Inclusive fitness is more generalized than just genetically related kinship, friendship, as a mutually supportive environment, arguably is the

altruism provided by individuals to their genetically similar others.

Rushton extended Hamilton's theory of inclusive fitness, and developed genetic similarity theory(J P Rushton, Russell, & Wells, 1985; J Philippe Rushton, 1989, 2005) . According to Rushton, "if a gene can better ensure its own survival by acting so to bring about the reproduction of family members with whom it shares copies, then it can also do so by bringing about the reproduction of organism in which copies can be found. This would be an alternative way for genes to propagate themselves. Rather than merely protecting kin at the expense of strangers, if organisms could identify genetically similar organisms, they could exhibit altruism toward those strangers as well as toward kin", he further explained that organisms identify genetically similar organisms by kin recognition, "a gene has two effects: it causes individuals who have it (1) to grow a green beard and (2) to behave altruistically toward green bearded individuals. The green beard serves as a recognition cue for the altruism gene. Altruism could therefor occur without the need for the individuals to be directly related." Thus, "genetically similar people tend to seek one another out and to provide mutually supportive environments such as marriage, friendship and social groups" (J Philippe Rushton, 1989). Although Rushton's genetic similarity theory is controversial and criticized (Allen, Eriksson, Fellman, Parisi, & Vandenberg, 1992; Cain & Vanderwolf, 1990), it still sheds some light on evolutionary perspectives on genetic homophily.

According to evolutionary biology, the preference to similar others are encoded in our genes. Though it is unlikely that people would directly observe genotypes of others around them, they would observe similar phenotypes, which suggest underlying genetic similarity. Preference and favorable acts to phenotypically similar others provide evolutionary advantages for genes to survive and thrive.

3.3.1.2 Active Gene-Environment Correlation

Another view on roles of genes in shaping friendship process involves active gene-environment correlation. This perspective has its origin from behavioral genetics which long has its interest in individual's ability to shape, structure and select environments. Evidence has shown the roles of genes in shaping friendship ties - a crucial aspect of social environment (Daniels & Plomin, 1985; Guo, 2006). Studies also suggested that genetic factors might influence individual's exposure to friends. Rowe and Osgood (Rowe & Osgood, 1984) used adolescent twins to study delinquency, their results revealed not only the anti-social behavior was largely heritable, but the correlation between the delinquency of an individual and the delinquency of his friends was genetically mediated, that is, the adolescents genetically predisposed to delinquency were also more inclined to seek each other out for friendship. A more recent study focused on adolescent alcohol and tobacco use (Harden, Hill, Turkheimer, & Emery, 2008) and found that some genetic factors related to an adolescent's own substance use and other genetic factors independent of substance use were both linked to increased exposure to best friends with heavy substance use.

Behavioral genetic theories have asserted an individual's genotype may be partially responsible for shaping, structuring and selecting environments that are compatible with his/her genetic disposition and allow optimum gene expression (Scarr & McCartney, 1983), the process by which genetic predisposition affects the environment one exposes to is known as active gene-environment correlation (Rutter & Silberg, 2002).

Active gene-environment correlation offers an important explanation for why people associate with genetically similar others. Many personality factors are at least partially heritable, and many behavioral traits, such as aggressive and violent behaviors are found with

genetic origin (Guo, Roettger & Cai, 2008). For example, for some adolescents, especially those with a proclivity to engage in delinquency, delinquent peer groups may be particularly appealing. Other teenagers, particularly those who are not inclined to delinquency, may veer away from delinquent peers and select more pro-social peers to be friend with. It is ubiquitous that individuals seek out and select friends that are compatible with their personalities, attitudes and behaviors. In this case, the genes that influence such personality, attitudinal and behavioral characteristics are also the genes that aid in the creation of social ties with others of similar genotypes. Genetically similar peers constitute the important social environment that is conducive to individual's own genetic expression.

3.3.2 Induced homophily

According to choice homophily perspective, genes may play active roles in shaping individual preferences for friends and thus lead to genetic homophily in social networks. Meanwhile, no behavior would occur in a vacuum, the second perspective – induced homophily – compliments the first perspective and emphasizes on the influences of social structures. Social structures may influence genetic homophily in the way that sort homogeneous people into socially and geographically proximate positions, thus shape potential pools of friends from which individuals may choose.

It is known that individuals are not uniformly distributed in geographic or social spaces; they are sorted into social structures by age, gender, race, occupation, education, wealth, workplace, neighborhoods, voluntary organizations etc. People in the shared structural environments are more homogeneous than the population at large. An individual's choice of relations is heavily influenced by the environments of his or her life that expose him or her to the similar acquaintances who share the environment, while effectively

excluding many different others (Feld 1981, 1982; Ibarra 1993). Thus, even without individual preference for similar others, the structural constrain on available choices and homogeneity of these choices will generate pattern of homophily. The pattern caused by homogeneity of structural opportunities for interactions is induced homophily (McPherson and Smith-Lovin 1987).

Not all structural factors that give rise to induced homophily would lead to genetic homogeneity. For example, friendship is known to sort upon age, while age homophily would not contribute to genetic correlations as people's genotypes do not vary by age. Only those upon which individuals are sorted but at the same time linked to genes would lead to correlated genotypes among friends. Geographic and social propinquities are the two major structural factors of this kind that lead to passive genetic correlations, which means friend selections are not based on certain genes, but genetic homophily based on these genes still arise as an outcome.

3.3.2.1 Geographic Proximity

The most basic structural source of genetic homophily is geographic propinquity. Due to human migration history and populations' adaptation to local environments, groups that are geographically distant tend to develop distinct frequencies of genotypes (Novembre, 2008), that is population stratification. Human migrated out of Africa from 100,000 years ago; subgroups of the origin population migrated to different geographic locations by random and possessed distinct subsets of genotypes from the origin population. Effects of the random changes in gene pool are maintained by assortative mating within geographically separated areas (Cavalli-Sforza, Menozzi, & Piazza, 1996). Genetic adaptations to local environments also added to distinctions in genotypes across geographical populations.

Though, it should be noted that 85%-90% of genetic diversity occurs between individuals within the population rather than between populations or continents, which was found in early studies on human diversity (Lewontin, 1972) and also confirmed by more recent molecular genetic evidence (Rosenberg & et, 2002). Moreover, most of these genetic variations across populations resulted from the migration history tend to be selectively neutral, that is, the genotypes tend not to confer functions that increase or decrease evolutionary fitness (Cavalli-Sforza, Menozzi, & Piazza, 1996). The neutral selection suggests that most of genetic differences across human populations are random and unrelated to human phenotypes.

Due to population stratification, genotypes may appear to be more correlated within geographically separate groups than between them. As we have more opportunities to contact with those who are close to us in geographic location than those who are distant, genetic homophily may appear simply as a byproduct of the geographic constrain.

An important social aspect associated with population stratification is race. Due to different origins, a small portion of genotypes are specific to each racial group, these genotypes are neutral and don't yield any selective advantages or disadvantages. Interpersonal interactions are more likely to occur within racial groups because of social or cultural reasons not because of individual preference for race related genotypes. Even though such genotypes do not have anything to do with friendship process, they still exhibit associations among friends as an outcome.

3.3.2.2 Social Proximity

Besides geographic locations, individuals are sorted into socially proximate positions. People vary genetically in disposition, they may actively choose occupations, workplaces,

neighborhoods or voluntary organizations they find conducive to their dispositions. Individuals may also be chosen or selected into different social contexts by institutional mechanisms out of their control. For example, the hiring process in some occupations may select for people with specific career orientation and skills, which might be associated with genetic predisposition. Similarly, individuals who fail to present adequate motivations and skills may drop out of the institutions. Genetic predispositions may cause individuals to enter or be selected into stratified social positions, within the social spaces, similar people gather. Even thereafter they associate randomly, genetic homophily may arise due to the social structures.

3.4 Hypotheses

Both choice homophily and induced homophily, which correspond to active and passive roles of genes, may give rise to genetic homophily among friends. To build up a baseline, the study first tests the following hypothesis:

Hypothesis 1: genetic correlation among friends (either reciprocal or non-reciprocal) should exceed the level expected by chance based on tructural factors (e.g. population stratification, age, gender and race).

Beyond the baseline, it is also important to examine whether social contexts which shape friendship opportunities would enable or eliminate the possibility of genetic homophily. The study then tests whether genetic homophily is contingent upon social contexts which structure friendship opportunities.

There are reasons to expect that genetic homophily will depend on friendship opportunities in social contexts. In highly constrained social settings, individuals barely have any choices regarding who they associate with, genetic homophily will be reduced or non-

existent because there is little room for individual preference to take place. On the contrary, if the environment provides great flexibility for individuals to choose their associates, the autonomy over friendship selection will lead to higher expression of genetic propensity, thus higher level of genetic homophily.

This leads to the second hypothesis:

Hypothesis 2: genetic homophily between friend dyads is contingent upon social constrain of friendship choice. Greater the constrain, weaker the genetic homophily.

Friendship formation is an initiation-response process which always starts with one individual making an offer of friendship to another, while the other person may or may not respond (Hallinan, 1978). It is reasonable to believe that friendships that are reciprocated are substantially different from those that are not (Hartup, 1996), A further question the study asks is whether such friendship dynamics give rise to differentiated levels of genetic similarities among reciprocal and non-reciprocal friends.

Actually the initial homophily is further strengthened by friendship dynamics, that is, friendships form and dissolve, overtime, ties between dissimilar individuals dissolve at a higher rate, which sets the stage for the even more similar individuals within the social space (Noel & Nyhan, 2011; O'Malley & Christakis, 2011). Reciprocal friendship is the type of social tie that lasts much longer than non-reciprocal friendship, which leads to the following hypothesis that

Hypothesis 3: genetic correlation among reciprocal friends should be stronger than that among non-reciprocal friends.

3.5 Methods

3.5.1 Data

The study uses two independent datasets with information about genes, socio-demographic attributes of respondents, as well as that of their social ties. The first data is from the National Longitudinal Study of Adolescent Health (Addhealth), a nationally representative sample of U.S. adolescents. The in-school survey and the first wave of in home survey were conducted in 1994, two years later in 1996, a second wave of in home survey followed. In all the three surveys, respondents were asked to name friends[6] and to provide more specific information about their friends. In addition to the friendship information, saliva DNA was collected for a subsample of siblings pairs (n>2000)[7]. These pairs include all adolescents who were identified as twin pairs, siblings, half-siblings or unrelated siblings raised together based on a screening of the in-school sample. The genetically informative subsample, which is composed of 2281 Addhealth respondents, is similar in demographic composition to Addhealth full sample (Jacobson & Rowe, 1998). The DNA sample is analytical sample of the study.

The study realizes limitations of the genetically informative sample, as the friendships analyzed here are only that between twins and full siblings, these friendships represent a small subset of the total friendships involving Add Health respondents. However, after carefully excluding friend pairs who are related, there is not much ground to suspect that the

[6] Friend nomination settings are different for saturated schools and unsaturated schools. In the 2 large and 10 small saturated schools, where all students from the in-school sample were surveyed in the in-home waves, students were instructed to nominate up to 5 male and 5 female friends. In the unsaturated schools, where not everyone in in-school sample was surveyed in the following in-home waves, each student was only asked to name 1 male and 1 female friends. The study acknowledges that friends nominated in saturated and unsaturated settings are different regarding the number and order of friends, but as the study interests in comparison between friends and random pairs, reciprocal and non-reciprocal friends, treating friends with different orders as equal would not bias the results.

[7] The DNA data is from Professor Guang Guo's project Illumina 1536 genotyping in Add Health (NSF's Human and Social Dynamics program BCS-0826913). Saliva DNA of the genetically informative subsample was genotyped using the Illumina 1536 array including a panel of 186 AIMs, and the genotyping yielded good quality data.

friendship subsample biases from the friendship full sample.

The study acknowledges that to study active roles of genes in friendship selection, a better question to ask would be whether genetically similar people are more likely to become friends, compared to random pairs. To answer this question, social network models such as exponential-family random graph models which estimate the probability of friendship ties as a function of whole network dynamics are required. While constrained by the DNA sample, social ties based on the subsample by no means can reflect the whole network process. Thus, the following analysis actually answers to the question: given that two people are friends, whether they are more genetically similar than random pairs.

The second data is from Roommate Study, a genetically informative study based on the campus of a southern university. Initially respondents were sampled based on university housing application. Given consideration of gender, smoking status, and other housing preferences (such as dormitory type, campus location etc.), housing applicants were randomly paired up with roommates. 2080 of the targeted respondents participated in online survey and provided saliva DNA sample (Guo et al, 2009). The full Roommate Study sample contains genetic information, while friendship information was gathered very differently in comparison to Add Health. In online survey, respondents were not asked to nominate their friends, instead, a question about their relationships with roommates was asked: whether [the roommate] was the type of person you would like to be friends with even if you were not roommates? Thus in Roommate Study, friends were only among roommates, who were randomly paired up in the first place, and may or may not become friends as they got along.

Saliva DNA was collected in both studies. The same Illumina GoldGate assay was used to genotype both saliva samples. After SNP screening and quality control, the two data

yielded an identical set of 100 SNPs from 21 behavior related genes, all the SNPs are of good quality with both individual missing and SNP missing less than 5% and Hardy-Weinberg disequilibrium test no more significant than 5%.

It would be ideal if friend similarity could be examined using the genome-wide data, however, given this set of 100 SNPs, the study may still yield meaningful results given that (1) the study design is very unique regarding taking advantage of the same panel of DNA collected in two very different friendship settings (2) the 21 genes chosen for genotyping are those that were already known to have significant effects on human social behaviors, thus are more targeted than genome-wide data and may yield significant results even given the limited number of SNPs .

3.5.2 Model

3.5.2.1 Estimating the Whole Set of SNPs Simultaneously

This study looks at genetic similarity between friends not on the basis of individual SNPs, but of all the 100 SNPs as a whole. The model fits all 100 SNPs simultaneously and estimates variation explained by all SNPs together. The following model estimates genetic similarity over all 100 SNPs for a specific friend pair.

$$Y = X\beta + Zu + \varepsilon$$

The outcome variable Y_{ij} is the dosage of minor alleles for a SNP, which is standardized with mean =0 and standard deviation=1. i=1,2 indexing two people in a friend pair and j=1, 2, …., 100 indexing the SNPs. X is a matrix of observed variables such as structural factors used as controls. β is a vector of the effects of X. Z is a matrix containing information of groups, SNP ID is used to specify groups when estimating friend-pair-specific correlation over all the SNPs. u is a random vector with $u_i \sim N(0, \sigma_u^2)$ where i=1, 2,…, 100 for

the number of SNPs. ε is a vector of residual effects with $\varepsilon \sim N(0, I\sigma_\varepsilon^2)$.

The model reports rho value which is the intra-class correlation and allows both positive and negative correlations. For a specific friend pair, greater the rho value, greater the similarity between friends over the 100 SNPs.

3.5.2.2 Simulation

The above model estimates overall genetic correlation based on the set of 100 behavior related SNPs for any given pair. To answer the question whether there is more clustering of genotypes among friend pairs than might be expected due to chance, the level of genetic clustering for friend pairs will be estimated, furthermore a simulation will be conducted to assess the level of genetic clustering by chance alone. Preserving the same origin population and sample size of friend pairs, a sample of random pairs will be generated for thousands of times in order to establish the baseline for genetic clustering –distribution of average rho values -- among random pairs.

This simulation test thus provides a way to test the null hypothesis that the observed genetic clustering among friends is no greater than the genetic clustering by random. Observed genetic correlation in friend pairs will be compared to the distribution of random genetic correlations, a low p value would suggest the observed friend genotype correlation was very unlikely to emerge due to chance alone.

3.5.2.3 Model Validation

To evaluate validity of this method for estimating genetic homophily, the study performs the analysis upon samples of known relatedness. Addhealth gathered 556 pairs of full siblings, 19 pairs of identical twins and 186 pairs of fraternal twins with both DNA and survey information. If the estimated genetic similarity for full siblings and fraternal twins is

around 0.5 and that for identical twins is around 1, the model yields reliable estimates on genetic similarity.

3.5.2.4 Control for Population Stratification

As noted, population stratification could result in correlated genotypes between two individuals who are simply geographically close, to reduce the likelihood of false-positives, the study uses principle component approach(PCA), which is widely applied in genetic studies and has been proven effective in controlling for population admixture (Price et al., 2006), to control for the population stratification.

The principle component analysis was implemented with the software EIGENSTRAT, the estimation yields three principle components for each genetic sample, and the principle components could be interpreted as bio-ancestry scores, which correspond to the percentages of Asian, European and African ancestry respectively, and the three scores add up to 100%. The bio-ancestry scores are introduced as controls in the above mixed model.

The estimation of population substructure relies on a panel of 186 ancestral informative markers (AIMs), which are selectively neutral genetic polymorphisms whose allele frequencies differ significantly across populations (Frudakis et al., 2003). The panel was specially designed for detecting continental populations of Europeans, Sub-Saharan Africans and East Asians, and studies have shown that regarding accuracy and informativeness of the AIM panel, 100 – 160 AIMs were sufficient when the sample size was roughly 1000 (Rosenberg et al., 2002). It is worth noticing that the AIMs panel is independent from the panel of behavior related genes.

3.5.3 Measures

3.5.3.1 Friend Pairs: Reciprocal vs. Non-reciprocal

In Addhealth, respondents were asked to nominate male and female friends respectively.

In roommate study, respondents were asked whether the roommate was the type of person he or she would be friends with even if they were not roommates. Respondents sometimes name family members as “friends”, in the analysis, all nominations to parents, siblings, half-siblings, adopted siblings, aunts, uncles, nephews, nieces and cousins are excluded out of friendship ties. Though in the two studies, measures for friend pairs are different, the way to distinguish reciprocal and non-reciprocal friend pairs is generally applicable: if the friendship nominations between two people are mutual, then the two people are considered as a reciprocal friend pair. If the nomination is unilateral, then the two people connected by the unilateral tie are considered a non-reciprocal pair.

3.5.3.2 Control Variables

The structural factors including self-reported race, gender, age and mother’s education, which may shape friendship process, are controlled when estimating genetic homophily. Race is measured by a categorical variable with mutually exclusive racial categories: Hispanic, multiracial, Non-Hispanic white, Non-Hispanic African America, Non-Hispanic Asian, and Non-Hispanic other. Gender is a dichotomous variable with male as the reference category, age is continuous and mother’s education is measured by a categorical variable (less than high school, high school, some college and college graduates and more than college) which captures mother’s highest level of educational attainment.

3.6 Results

The study first presents descriptive statistics for AddHealth and Roommate Study analytical samples. By comparing AddHealth DNA sample with the wave III full sample, it

shows that the analytical sample is comparable to AddHealth full sample with regard to all social demographic variables. The table also presents a comparison between AddHealth and Roommate Study analytical samples, Roommate Study sample was composed of college students who were older than adolescents in AddHealth sample. Roommate Study sample had more whites and females, and had parents with higher level of education. Though variation in socio-demographic backgrounds may lead to different friendship processes among AddHealth and Roommate Study respondents, the study would argue that the major difference came from contextual constrains on friendship opportunities. For AddHealth, respondents had the opportunity to name any other student in the school as a friend; while for Roommate Study, the friendship opportunity was constrained between randomly paired roommates. The contextual constrains, rather than socio-demographic backgrounds, may shape genetic homophily among friends.

Table 10 presents number of friend pairs from the two samples, there are 895 pairs of friends out of 2281 AddHealth DNA sample, less than ¼ of friends are reciprocal pairs. Out of 2080 Roommate Study sample, there are 426 friend pairs and around half of them are reciprocal.

The study first answers to the question whether genetic correlation among friends exceeds the level expected by chance based on structural factors. As shown in table 3, for 206 pairs of AddHealth reciprocal friends, the average genetic correlation based on 100 behavior related SNPs, presented by average rho value, is 0.047, even after controlling for bio-ancestries and other socio-demographic covariates. For 206 random pairs (related pairs were excluded) which were drawn from the same AddHealth DNA sample, the average genetic correlations fell into a distribution with 0.0023 mean and 0.0074 standard deviation. Average

genetic correlation for 206 AddHealth reciprocal friends is significantly greater (p=1.12E-11) than the level of random pairs, thus the first hypothesis is supported. The hypothesis is further validated by AddHealth non-reciprocal friends, though the genetic correlation is slightly lower than that for reciprocal friends (0.039 vs. 0.047), it still exceeds the level expected by chance (p=1.02E-6).

For Roommate Study, the results tell a very different story. The average genetic correlations for reciprocal and non-reciprocal friends are 0.0075 and 0.0048 respectively, neither exceeds the level of random pairs. On one hand, the result is not surprising regarding that Roommate Study friends were randomly paired roommates in the first place, thus their genetic similarity level should not exceed the level of random pairs. On the other hand, it supports the second hypothesis and unravels the importance of contextual constrains in shaping friend genetic homophily.

Table 11 also shows that average genetic similarity for full siblings and DZ twins is around 0.5 and that for MZ twins is around 1, which validates the approach for estimating genetic homophily.

Figure 3 presents comparison of observed rho value for various pairs to simulated rho values in 3000 simulations in which pairs were randomly drawn, keeping the original population and sample size of observed pairs. The results show that for AddHealth friends, both reciprocal and non-reciprocal, observed rho values fall in the tails of the simulated distribution and therefore are very unlikely due to chance. Observed rho values for Roommate Study friends are less extreme, and therefore more likely to have occurred by chance.

Table 12 answers to the question whether reciprocal friends are more genetically alike

than non-reciprocal friends. According to two sample mean test, the average genetic similarity for reciprocal pairs is moderately greater than that for non-reciprocal pairs ($p=0.069$), and figure 4 visualizes the comparison between the two types of friend pairs regarding pair specific genetic correlations. To check robustness of the findings, the comparison was conducted for in-school, wave I and II pairs respectively, and the results show that the genetic similarity for reciprocal friends appears to be consistently greater, though at a moderate level, than that for non-reciprocal friends across waves.

3.7 Conclusion and Discussion

The results are both consistent and provide important extensions to prior work. First, with an expanded understanding of the relationship between genes and the behavioral outcomes, the results show that, instead of individual polymorphism, friends are more alike than random pairs based on a set of behavior related genes. The finding is important as it provides a more comprehensive baseline for friendship genetic homophily, and moreover it suggests that friends might be positively correlated on some markers and negatively correlated on others, while overall it presents genetic homophily instead of heterophily.

Beyond the baseline genetic homophily, the finding that reciprocal friends are moderately more alike than non-reciprocal friends further supports the role of genes in friendship formation. As friendship is an initiation - response process, when a friend nomination is received, whether to reciprocate the relationship involves an additional round of friendship decision. Greater genetic similarity among reciprocal friends does suggest that genetic homophily is driven by individual preferences for genetically similar others as friends.

Inclusive fitness/genetic similarity theory and active gene-environment correlation provide theoretical explanations for why genetic similarity may drive the way people make

friends, besides the interests in exploring genetic factors which explain complex behaviors such as friendship, the study also reveals the important roles played by social structures in the extent to which genetic factors matter. Social factors may structure relationships that enable or eliminate the possibility of genetic homophily in friendships.

The level of genetic homophily varies considerably across two studies, and differences in contextual constrains on friendship choice account for observed differences in genetic homophily. AddHealth allowed respondents to nominate friends freely, while Roommate study asked whether respondents considered the randomly assigned roommate as a friend. Differences in the two settings have great implications for expression of friendship preference. In Roommate study, even roommate become friends eventually, they were randomly paired up in the first place, thus there had been no room for selecting preferred friends. For Addhealth, friendship formation occurs in a more complicated setting (1) Addhealth provides big enough a pool of potential friends to choose from (2) Addhealth also offers foci of activities (classes, interest clubs etc.) which bring together disproportionately homogenous sets of people as well as the possibility of friendship transitivity.

The study focuses on two school settings while the findings are relevant in more broadly defined social environments. In highly constrained social settings, the influence of genetic factors on friendship formation will be reduced or nonexistent because there is very little room for subtle genetic influences to take place. While in environments which provide greater flexibility and autonomy for individuals to select their friends, genotype may become a relatively more important factor with respect to friendship selection. The study highlights the importance of both active process of friend selection and contextual mechanisms for genetic homophily among friends, furthermore the importance of incorporating contextual

factors while exploring genetic explanations for behaviors such as friendship, also the necessity of integrating genetic theories and variables into the exploration of sociological inquires.

Finally, the study asks for further exploration of this line of inquiry. First, it would be interesting to look over the whole genome and to explore the homophily with broadened view of genome. Second, the study asks for a more comprehensive framework which incorporates phenotype measures in the current model. Lastly, the results about reciprocal and non-reciprocal friends are tentative and ask for replication and further investigation.

CHAPTER 4

FRIENDSHIP RECIPROCITY AND SOCIAL NETWORK CENTRALITY: HOMOPHILY OR STATUS ASYMMETRY

4.1 Introduction

Reciprocity is one of the expectations about affective relations (Laursen, 1993). When it comes to friendship between two individuals, the relationship is commonly assumed to be mutual in nature (Freeman, 1992), that it, if Alice sees Bob as a friend, people tend to believe that Bob sees Alice as a friend as well. While, as suggested by a study as early in 1970s, friendship is an initiation-response phenomenon, the process of making friends always starts with one individual making an offer of friendship to another, while the other person may or may not respond (Hallinan, 1978). Actually, non-reciprocated relationships have been rather common in friendship networks, for example, only 40% of the friendship nominations in adolescent social networks were reciprocated (Mouw & Entwisle, 2006). Similar phenomenon was observed in adult social networks as well, another study (Mollica, Gray, & Trevino, 2003) found that over one third of the friendship ties they initially measured were unreciprocated. Unreciprocated friendships tend to lead to lower levels of positive feelings than mutual friendships (Mendelson & Kay, 2003), thus unilateral ties are rather fragile in nature. An offer of friendship can make a mutual tie more likely to arise relative to nonexistent relationship, while unilateral friendships are fundamentally short lived if they fail to become reciprocal (Rivera et al., 2010).

It is reasonable to believe that friendships that are reciprocated are substantially

different from those that are not (Hartup, 1996). In James Coleman's seminal essay on social capital, reciprocity of expectations and norms is one important feature of social capital; similarly, reciprocal friendships are superior social capital resources than non-reciprocal ones; reciprocal friends are also more emotionally supportive and better sources of information gathering (Plickert, Côté, & Wellman, 2007). For adolescents, reciprocal friendship is associated with increased school belonging and better academic performance (Vaquera & Kao, 2008).

Friendship reciprocity is important not only because of its impact on individual outcomes, but also because of its implications for social network properties. Schaefer (Schaefer, 2012) used a laboratory experiment and revealed that overtime individuals adjusted their friend seeking strategies to increase reciprocity, more interestingly, he found that non-reciprocity was a mechanism resulting in increasing level of homophily in the network.

Despite the importance of friendship reciprocity, it is still a relatively open area of inquiry (Rivera et al., 2010). A fundamental question which has not been extensively studied is why some relationships are reciprocated while others are not. Unreciprocated relationships always suggest status hierarchies, Gould (Gould, 2002) commented "someone who pays less attention to you than you pay to her implicitly asserts that she is superior to you in status. If you do not respond by withdrawing your attention, you have implicitly agreed." Similarly, the study of Vaquera and Kao (Vaquera & Kao, 2008) found that adolescents from higher SES families, measured by mother's education, enjoyed higher rates of friendship reciprocity. Olk and Gibbons (Olk & Gibbons, 2010) explored friendship reciprocity among executive MBA students and also found that unreciprocated friendship flew from a person with lower

gatekeeping power -- the access and control over information traveling through the local grapevine -- to a person with high gatekeeping power. All the studies suggest that status may play an important role in friendship reciprocation.

Social status has many dimensions which are difficult to measure comprehensively. This study primarily focuses on one status dimension regarding one's position within the social network -- the social network centrality. Centrality is a fundamental measure of the importance of individuals in social networks; importance of this measure was already stressed in early works of Bavelas (Bavelas, 1948) and more recent work such as Wasserman and Faust (Wasserman & Faust, 1994). Centrality in social networks is increasingly used as explanations of individual outcome and decisions. Empirical evidence has shown that centrality was important regarding peer influence (Dana L Haynie, 2001), creativity of workers (Perry-Smith & Shalley, 2003) and the flow of information (Borgatti, 2005). Due to greater impacts on others in the network, better access to information, human capital and other precious resources, central position in a social network always equates to power and status.

The study looks at centrality in global instead of local networks. Global network centrality concerns not only direct connections to individuals, which compose local or ego networks, but also indirect connections through friends of their friends, friends of their friends' friends etc., which determine the position that an individual occupies in the social space. Although global network centrality could be an ideal proxy for one's relative status in the peer group, this measure also presents challenges for the study, such as reserve causality. The study interests in the causal effect of one's network position on his or her friendship experience, while it is equally plausible that one's friendship actions may have consequences

for the space he or she possesses in the social network as well. One approach to deal with the reserve causality is to set up appropriate temporal ordering. Fortunately, the National Longitudinal Study of Adolescent Health (Add Health) measured friendship and network dynamics in multiple waves and provided valuable opportunities to study the causal effect of social network centrality at an earlier time point on the consequent friendship reciprocity.

Furthermore, although individual decisions on forming and declining direct friendship ties may have influences on his or her network position, the influences are likely to occur within the boundary of ego networks, and it is reasonable to believe that the effects that ripple through local networks to global networks are limited, given the nature of the global network structure that is composed of a few direct ties but numerous indirect ties which are far beyond control of individuals. Additionally, especially for children and adolescents, many of them have to wait until a peer group selects them, although they may later decide to accept or not. The number and type of networks open to them are limited by many factors including age, gender, culture, and socio-economic background, which may be even more important factors than personality or behavior orientation (Brown, Lohr, & Trujillo, 1990). This suggests that individuals do not have as much control and freedom in deciding their positions in the global network.

By examining how network positions of respondents and people who nominated them as friends affect the likelihood of friendship reciprocity, the study intends to answer the basic question: when it comes to the decision of building a mutual relationship, what do people value about the potential friends? The study is conducted within the context of school-based networks using the National Longitudinal Study of Adolescent Health (AddHealth). Although not all friendships of adolescents occur within schools, schoolmates compose the majority of

friend pool adolescents are likely to select, thus school setting allows to capture much of adolescents' social universe. Understanding adolescent school-based networks is of great importance because adolescence is a life stage featured by increased influence from peers and decreased influence from parents (Larson & Verma, 1999), and adolescents learn social rules, negotiate social roles and gain social status largely through interacting with peers. Furthermore, the specific school setting allows us to study the general properties of friendship reciprocity that likely generalize to other social contexts and broader populations.

4.2 Theories

The study draws on two theoretical orientations to explain friendship reciprocity: homophily and status asymmetry. Both perspectives concern about a basic question, what do people prefer or value about their potential friends? Do people prefer someone similar to them or someone with higher status as friends? In brief, the homophily perspective emphasizes on individual preference for similarity and infers that similarity enhances the likelihood of relationship reciprocity. Status asymmetry perspective postulates that when individuals share common definition of status, they prefer higher status partners, thus higher status is always associated with favorable friendship outcomes including reciprocity.

4.2.1 Homohpily

Several theories explain to whom people make friends, the most prominent one is homophily perspective: birds of a feather flock together. Here the term "homophily" refers to "the love of the similar", which is a tie formation mechanism based on individual preference. In the second paper "genetic homophily in social networks", the term was used slightly differently to connotate compositional features of networks which are outcomes of such tie formation mechanisms, which is a synonym for "homogeneity".

This line of study can be traced back as early as 1950s, Lazersfeld and Merton's (Lazersfeld & Merton, 1954) analyzed friendship selection among adults and used the term "homophily" to refer to "a tendency for friendships to form between those who are alike in some designated respect" . This landmark work gave birth to a rich tradition of research that investigated the tendency toward similarity in a variety of social settings, and it is universally found that the greater the similarity between two individuals, the more likely they are to establish a connection (Mcpherson, Smith-lovin, & Cook, 2001).

People may prefer similar others for many reasons, in general, homophilous relationships are more rewarding: homophily has been found to reduce interpersonal strain (Newcomb, 1961), facilitate communication (Rogers & Bhowmik, 1970), reinforce beliefs and identity (Byrne, 1971) and increase coordination (Cole & Teboul, 2004).Similar others also have longer (Ledbetter, Griffin, & Sparks, 2007) and more intimate friendships (Biesanz, West, & Millevoi, 2007; Heller & Wood, 1998).

Beyond individual preferences, there are structural reasons that may give rise to friendship reciprocity as well. It is intuitive that people in socially and physically proximate positions have greater opportunities to interact with and more ease to maintain the relationship with each other than people who are far apart. Familiarity and interaction are simply the reasons that cultivate mutual friendships.

Thus due to individual preferences or structural factors, the homophily perspective predicts that the more similar the ego and alter are regarding social status, the more likely the friendship is reciprocated.

4.2.2 Status Asymmetry

Homogeneity regarding numerous types of attributes among friends seems the most

robust empirical regularity in social world (Mcpherson et al., 2001), while the resemblance as an outcome may or may not reflect the true preference in intimate relationships in the first place.

The role of attractiveness in mate selection first attracted researchers' interest (Roth & Sotomayor, 1990). It is widely observed that couples often have similar levels of attractiveness: do people prefer similarly attractive partners in the first place? Opposite to the matching hypothesis which was proposed in 1960s (Walster, Aronson, Abrahams, & Rottman, 1966) and predicted that individuals prefer partners who are as attractive as they are; early studies found that most subjects preferred partners who are more attractive than themselves. Another example is depression homophily. Depressed youth are friends with each other not because they prefer each other, but they are more passive and withdrawn, their peers do not select them as friends, thus they are left with only themselves to befriend (Schaefer, 2012).Deviations from preference for similarity may also be observed for other less desirable traits such as aggression (Sijtsema et al., 2009) and obesity (Crosnoe, Frank, & Mueller, 2008). Rejection, exclusion and avoidance rather than preference for similarity were used to explain the homogenous ties. Thus the status asymmetry perspective assumes that when individuals share a common definition of value, they always prefer higher value partners, the differences between ego and alter in value result in status asymmetry.

However, the status asymmetry perspective may not need to be seen as contradicting the homophily perspective. In the context of extending friendship offers, one may seek for higher value partners in the first place, while as predicted by social exchange theory (Emerson, 1976), the gesture may not be rewarded in return, the fear for rejection or the desire for mutuality drives the individuals to adjust their preferences and results in a shift

from higher to similar value targets who are more likely to name them as friends in return. In the context of friendship reciprocation, given that the friendship offer is already extended from ego to alter, whether alter responds to the gesture depends on alter's preference and his or her evaluation on ego's value, while such preference is not subject to adjustment due to fear for rejection. Moreover, the two perspectives actually complement each other by looking at both distance and directionality in the dyadic relations. The homophily perspective concerns about the distance or difference between ego and alter regarding attributes of interest, while status asymmetry perspective adds up another dimension by looking at the relative positions of ego and alter in the value spectrum, and gives additional leverage to answer the question given the distance or similarity between ego and alter, whether their relative position matters.

4.3 Hypotheses

Based on the two theoretical orientations, the study examines how social network centrality influences friendship reciprocity through mechanisms at relational level. Relationship formation always requires that individuals find something attractive in one another to motivate further interaction (Fine, 1980). Similarly, when ego extends friendship to alter, whether the friendship will be reciprocated depends on ego's "attractiveness" to the alter. According to the homophily perspective, the alter would view someone who possesses a position that is similar or close to his or her own as attractive, thus smaller the centrality difference between ego and alter, greater the probability ego's nomination being reciprocated. This gives rise to the **homophily hypothesis**:

Hypothesis 1: the social network position distance between ego and alter matters, greater the distance, smaller the likelihood of reciprocation.

Homophily perspective only addresses the issue regarding network distance between two people linked by a friendship nomination; another equally important dimension involves the relative position of the two people, which is the focus of status asymmetry perspective.

Status asymmetry perspective predicts that the alter would view someone who possesses a higher status as attractive, thus the relative position of ego and alter matters, when the relationship is extended from someone in a more central position to the other in a less central position, it is more likely to be reciprocated than that extended from the reserve way. The study derives the following hypothesis based on **status asymmetry perspective**:

Hypothesis 2: friendships offered by people with a more central position in the network to people with a less central position are more likely to be reciprocated than friendships offered in the reserve way.

As acknowledged in above theoretical discussion, the two perspectives may not need to be viewed as contradicting each other; actually they may coexist and complement each other. That is, if the status asymmetry hypothesis is true, people in general prefer to reciprocate nominations from people with higher status, among all the nomination senders with higher status, people may still prefer those with smaller status distance, which supports homophily hypothesis as well. Meanwhile, if the homophily hypothesis is true, the likelihood of friendship reciprocity should decrease with status distance, but the negative effect of status distance may differ for dyads where the nomination sender is of higher status and where the nomination sender is of lower status. If the negative effect of status distance is smaller when the nomination sender is of higher status, it supports status asymmetry hypothesis as well.

Hypothesis 3: if both the homophily and status asymmetry hypotheses are true, the negative effect of status distance on friendship reciprocity should be smaller when the

nomination sender is of higher status.

4.4 Data and Methods

4.4.1 Data

The study uses data from in-school and wave I of the National Longitudinal Study of Adolescent Health (Add Health), whose rich social network data provides a unique opportunity for investigating the associations between network centrality and friendship reciprocity in a nationally representative sample. Add Health uses a multistage, clustered sampling design which selected a nationally representative sample of 132 middle and high schools. Add Health was first initiated with the in-school survey between 1994 and 1995. A subsample of participants from the initial in-school survey was then selected for an in-depth in-home survey two months later.

As the study intends to rule out reverse causality by setting temporal ordering of network centrality and friendship reciprocity, in-school centrality will be used to predict wave I friendship outcomes. The two-month window between in-school and wave I provides a great time opportunity to study friendship reciprocity, first it allows one's social network position as an antecedent of following friendship decision, meanwhile, centrality measure at in-school survey serves as a good proxy for one's network position at wave I when friendship decision occurred, as one's network centrality should not change much during the two-month time window.

In both in-school and wave I in-home surveys, participants received a roster of all students enrolled in their school and their sister feeder school. At in-school survey, all participants were asked to nominate up to five male and five female friends. At wave I, only a subsample of in-school respondents was surveyed. In 2 large and 10 small saturated schools,

in which all in-school respondents were surveyed, respondents were asked to nominate up to five male and five female friends; in other unsaturated schools, where not every in-school respondent was surveyed, respondents were asked to nominate only 1 male and 1 female friends. The nominated friends who were not found on the roster are excluded from the analyses because no data were available for them. Friendship network measures were constructed based on the in-school friend nominations, which provided global pictures of school networks.

In the in-school survey, all students in attendance in the participating schools were given self-administered questionnaires, which yielded a sample of 90,118 adolescents. In principle, all students in schools were targeted which allowed generating complete school network data. While among the 132 participating schools, only 129 out of them had more than 50% of the student body completed the in-school questionnaires. In order to get accurate network measures, only the 129 instead of 132 schools were used in following analyses.

Within the 129 schools, 14,319 students participated in the wave I in-home data survey as well. 10,498 out of the 14,319 students had valid friend nominations and identifiable friends. Thus, the analytical sample is composed of the 10,498 adolescents with valid in-school network measures, wave I in-home survey information and friend nominations. Comparing with AddHealth full sample, the analytical sample has slightly greater proportions of females, whites and students with higher parental education, overall, it is comparable to wave I in-home survey sample regarding socio-demographic characteristics.

4.4.2 Measures

Dependent variable

As the study interests in the impacts of in-school social network position on wave I

friendship decision, the major outcome of interest – friendship reciprocity will be constructed based on wave I friend nominations. **Friendship reciprocity** concerns whether a friendship nomination from ego i to alter j is reciprocated, and is measured by a dummy variable with value 1 indicating a reciprocated nomination and value 0 indicating an unreciprocated one.

A potential concern associated with the reciprocity measure originates from the study design which instructed respondents to nominate up to a certain number of friends. Friendship nominations are always directed to popular individuals; meanwhile, the person who receives a lot of friendship nominations probably also possesses a central network position. Given the limited number of nominations one may reciprocate, the greater the number of received nominations, the smaller the proportion one may reciprocate, it may create the phenomenon that people in the center of social networks are also less likely to reciprocate other's nominations, which may just be the artifact of the friend nomination design.

Robustness test is conducted to address the potential bias, the test asks the question, contingent upon the friendship nominations one receives and the constrained number of nominations one may reciprocate, and how one would make the reciprocation decision. The question is examined using discrete choice model, in which received nominations are viewed as discrete alternatives and friendship reciprocity involves choices between the alternatives.

Independent variables

Drawing upon the centrality measures used by Haas et al (Haas, Schaefer, & Kornienko, 2010), the study measures one's position in a global network by **influence domain** and **Bonacich centrality**. The two measures are parallel to meanwhile different from local network measures in-degree and out-degree, regarding that in-degree and out-degree are

constructed upon direct connections, while global network measures including influence domain and Bonacich centrality are based upon both direct and indirect connections.

Influence domain and Bonacich centrality capture different aspects of centrality[8] and how active individuals are in the broader school context, not just with their groups of friends. Using outgoing ties, Bonacich centrality measures the influence of an individual in a network. It assigns relative scores to all individuals in the network based on the concept that connections to high-scoring individuals contribute more to the score of one's own than equal connections to low-scoring individuals. More specifically, it weights ego's centrality by the centrality of people to whom he or she sends nominations (Phillip Bonacich, 1987).

Bonacich centrality score would be high when individuals are connected to those who are themselves well-connected and is often used as an indicator of sociometric popularity (P Bonacich, 2007). It is calculated according to the following formula:

$$\text{Bonacich Centrality of } \mathrm{X}(\alpha, \beta) = \alpha * (\mathrm{I} - \beta * \mathrm{X})^{-1} X1$$

Where X contains all friendship nominations in the form of an adjacency matrix; α is a scaling factor; β is a power weight reflecting the degree of dependence of actor's importance on the extent of importance of the alters to whom the ego sends friendship ties; I is the identical matrix; and 1 is a vector of 1s.

Given that the survey design that only a certain number of friendship nominations were allowed, Bonacich's centrality for the gregarious people who sent out more than 10 friendship offers yields an underestimation. Examining centrality through incoming along with outgoing ties is important for overcoming such measurement errors. Influence domain is the other centrality measure which is based upon the number of incoming ties, others who nominate ego. Influence domain measures the number of alters who can reach ego through

[8] Pearson's correlation between influence domain and Bonacich centrality is 0.19 based on analytical sample of the study.

direct and indirect connections. Given that the measure is sensitive to the size of a school, it was standardized by dividing the size of school network to create a proportion.

On one hand, if results for the two measures of global network positions converge, it provides cross validation for measurement and findings of the study. On the other hand, both measures concern whether individuals are located in prominent position within their friendship networks according to whether they are involved with many others. It does not matter whether the involvement is due to being the recipient of many friendship ties or the source of the ties – what is more important is that a central individual is involved with many relationships with friends (Dana L Haynie, 2001). The centrality and influence scores are standardized in the following analyses.

Controls

By modeling controls, the study intends to take into account (1) factors that may influence both friendship reciprocity and social network centrality (2) mechanisms besides homophily and status asymmetry that may also lead to the influence of network position on friendship reciprocity, such as transitivity.

Local network variables

Out-degree [9] , which indicates gregariousness, is the number of friendship nominations that an adolescent makes (the unidentifiable friends who are beyond the school roster are also counted) . It has an upper boundary given that each respondent was only able to nominate up to a certain amount of friends according to the questionnaire design. **In-**

[9] Out-degree is local network counterpart for the global network measure centrality. Pearson's correlation between out-degree and centrality is 0.88 based on analytical sample of the study. Although the correlation is rather high, it is ok to model both variables in a model as multicollinearity should be of concern if the correlation between a pair of predictor variables is above 0.9 (Belsley, Kuh, and Welsch 1980).

degree[10] , which indicates popularity, is the number of friendship nominations that a focal adolescent receives, which is less affected by the questionnaire design. The two measures are based on ego network and may moderate the influence of network centrality on friendship reciprocity.

Other local network mechanisms may also lead to friendship reciprocity. For example, it is well known that friendships are always organized around cliques – the small group of people who interact with each other more regularly and intensely than others in the same setting. In the closely knit (dense) cliques, friendship nominations may be more likely to be reciprocated than is the case in less cohesive ones. **Ego network density** is a measure of such clique property. It is calculated by dividing the number of ties between ego and alters by the total number of possible ties between alters. This results in a proportion representing the level of interconnectedness among each respondent's friends. Ego network density should be positively associated with reciprocity as it provides greater opportunities for transitivity – shared friends are likely to foster new connections -- to occur.

Socio-demographic attributes

Socio-economic Status (SES): much research suggests that schools are more hospitable to higher SES than lower SES youth (Lareau, 2000). Furthermore, some argue that youth of higher SES background develop friendships with greater ease. Coming from an advantaged family may help some adolescents gain both social network centrality and reciprocity because they are more socially desirable in school. It is possible that more advantaged students become socially desirable because they present the desired life style: the latest gadgets, fashionable clothes, the nicest car etc. As Milner (Milner Jr., 2004) argues in his book, adolescents' status strives to consumerism. The variable to measure respondent's

[10] Pearson's correlation between in-degree and influence is 0.44.

socioeconomic background is a categorical variable (less than high school, high school, some college and college graduates and more than college) which captures the highest level of educational attainment of his or her parents.

Gender: it is well known that compared to males, females maintain stronger relationships and share higher levels of disclosure (Billy & Udry, 1985). On the other hand, compared to girls, boys tend to have a more open network that is less intimate, more volatile and more likely to include new friends over time (Belle, 1989). Thus we would expect that girl's relationships are more likely to be reciprocated compared with boys' friendships[11].

Race: race is measured by a categorical variable with mutually exclusive racial categories: Hispanic, multiracial, Non-Hispanic white, Non-Hispanic African America, Non-Hispanic Asian, and Non-Hispanic other. The study of Vaquera and Kao (Vaquera & Kao, 2008) on Add Health found that among all the racial groups, Asian Americans are most likely to have reciprocal friendships[12].

Age: age might be positively associated with status, thus the network centrality in adolescent networks. Moreover, the likelihood of friendship reciprocity may increase with age. Age is used as a continuous variable in the control[13].

Grade: It show that the average student shares six classes with students in the same grade and one elective class with students in all other grades (Zeng & Xie, 2008), as individual data were not available on tracking, grade is used as proxy for shared classes.

Gpa: GPA variable measures average of self-reported grades on Mathematics,

[11] In analyses, it doesn't exhibit significantly greater likelihood of friendship reciprocity among girls than boys after controlling for status distance and asymmetry between the potential friends (data not shown).
[12] No significant differences in friendship reciprocity regarding race were found after controlling for status distance and asymmetry (data not shown).

[13] Similar to gender and race, age doesn't seem influence likelihood of reciprocity after controlling for status distance and asymmetry (data not shown).

History/Social Studies, English/Language Arts, and Science, and ranges between 0 and 4.

PVT score: Add Health picture vocabulary test (AHPVT) has 78 items and is an abridged version of the Peabody Picture Vocabulary Test (PPVT) especially designed for Add Health. The PPVT estimates students' verbal ability or scholastic aptitude (Dunn, 1981), and has been widely used as a measure of academic performance (G. Duncan & Brooks-Gunn, 1997). The study uses wave I in-home test scores.

Relational characteristics

Opportunities for interactions, as well as the activities that the friends perform together, are important predictors of the strength and influence of the friendship (Dana L Haynie, 2001), thus influence the likelihood of friendship reciprocity. The study measures such relational characteristics by looking at number of types of activities one performed with a specific friend in previous week (ranges from 0 to 5), the five activities that one student can report to have done with friends include going to his/her house, seeing the friend after school, spending time together during the weekend, talking on the phone with the friend and talking to the friend about his or her problems(Vaquera & Kao, 2008).

School Characteristics

One important school characteristic involves school network property. Index of mutuality is a global network measure of friendship reciprocity. Katz and Powell (Katz & Powell, 1955) developed the mutuality index which measures the tendency for individuals in a group to reciprocate choices. The index is based on the expected number of mutual dyads given a random network with the same distribution of out-ties as that observed in the data(Wasserman & Faust, 1994). Greater the index, more likely a nomination in the network is reciprocated as well.

Other school characteristics which may condition the type of friendships formed include school size and urbanicity. Literature suggests that it is easier to establish more intimate friendships in small schools than big schools (Bryk, 1996). Mutual friendship, as a type of intimate relationship, is expected to be negatively correlated with school size. Similarly, suburban and rural schools are expected to promote reciprocal friendships as well due to their size and the greater intimacy of smaller communities.

4.4.3 Models

As friendship reciprocity is based upon already extended friendship ties, when it comes to modeling, the first step is to provide a running definition for friendship ties. There has not been much consistency regarding the way to measure friendship ties, borrowing from Vaquera and Kao (Vaquera & Kao, 2008), the study looks at first-list same-sex friend nominations only, and thus reciprocation is defined as the best same-sex friend reciprocates ego as best friend as well.

The study acknowledges that there are multiple ways to define friendship reciprocation, for example Vaquera and Kao (2008) provided another definition that the best friend reciprocates ego's nomination regardless the rank of friendship orders. Zeng and Xie (Zeng & Xie, 2008) used three ways to classify friend nominations: (1) best friend selection: ego nominates the single best friend (2) ordered selections: ego nominates up to a predetermined number of friends in order of closeness (3) unordered selections, ego nominates up to a predetermined number of friends without specifying the order of closeness, thus there are three ways to define friendship reciprocation accordingly. The current project decides on the above definition based on several considerations: first, both studies found that different definitions all yielded consistent results (Vaquera & Kao, 2008; Zeng & Xie, 2008);

second, adolescent friendships are fickle in nature, examining the relationships between same-sex best friends may help yield more stable results. The third consideration is based upon AddHealth research design, as students in saturated and non-saturated schools were asked to nominate different number of friends, to eliminate potential bias due to measurement, the study only looks at first-list same-sex friend nominations which were measured consistently in saturated and non-saturated schools.

The analysis is conducted with two models: the mixed effect model uses friendship nomination as analytical unit, and the discreet choice model uses individual as analytical unit.

For the mixed model, the likelihood of a nomination from ego i to alter j being reciprocated is a function of status distance and status asymmetry between i and j, local network and socio-demographic attributes for i and j, relational characteristics between i and j and school characteristics. Considering that some pairs of friends may enter the analysis twice (one for each nomination), standard regression methods which assume independency of observations may produce biased estimates of standard errors and result in increased probability of type-I error. Thus the study adjusts for clustering of observations by using mixed model with pair level random intercept.

As discussed earlier, the study also uses discreet choice model as a robustness check and answers to the question: among all egos who nominated alter as best friends, who would alter reciprocate as best friend as well. The model relates the choice made by alter to the attributes of both ego and alters, and estimates how alter's reciprocation decision will change under changes in attributes of egos. The discrete choice model takes the mixed logit form, which is highly flexible and can approximate any random utility model (McFadden & Train, 2000).

According to the classic book on discrete decision modeling by Train (Train, 2002), the mixed logit probability can be derived from utility maximization. In a straightforward specification, the decision maker faces choices among J alternatives. The utility of person n from alternative j is specified as

$$U_{nj} = \beta_n' x_{nj} + \varepsilon_{nj}$$

Where x_{nj} are observed variables that relate to the alternative and decision maker, β_n is a vector of coefficients of these variables for person n representing the person's attributes, and ε_{nj} is a random term that is i.i.d.

The coefficients vary over decision makers in the population with density $f(\beta)$. The density is a function of parameters θ that represent the mean and covariance of the βs in the population. β varies over decision makers rather than being fixed. The mixed logit probability of person n choosing alternative i is

$$P_{ni} = \int \left(\frac{e^{\beta' x_{ni}}}{\sum_j e^{\beta' x_{nj}}} \right) f(\beta) d\beta$$

$f(\beta)$ is usually specified to be normal: $\beta \sim N(b, W)$ with parameters b and W that are estimated. The integral for this choice probability does not have a closed form, so the probability is approximated by simulation.

The ratio of mixed logit probabilities $\frac{P_{ni}}{P_{nj}}$ depends on all the data, including attributes of alternatives other than i or j, thus mixed logit does not exhibit independence from irrelevant alternatives (IIA) as the conditional logit model. SAS procedure PROC MDC is used to implement the mixed logit model.

For either random effect model or discrete choice model, status distance and asymmetry are measured as following. The study draws upon Zeng and Xie (Zeng & Xie,

2008), for a nomination from i to j, σ_{ij} indicates status asymmetry between i and j.

σ_{ij}=1 if status i > status j; σ_{ij}=0 if otherwise.

|status i – status j| measures status distance between i and j.

To model status distance and asymmetry simultaneously, the analysis includes interactions between status distance variable and the indicator for status asymmetry:

$\beta_1 * \sigma_{ij} + \beta_2 (1 - \sigma_{ij})$ |status i – status j| + $\beta_3 * \sigma_{ij}$|status i – status j|

This specification allows the effect of status distance to differ for dyads where the nomination sender is of higher status (denoted by ego > alter) and where the nomination sender is of equal or lower status (ego <= alter).

If status asymmetry hypothesis is true, the fact that nomination sender is of higher status should have a positive effect on friendship reciprocity, thus β_1 should be positive and significant.

If homophily hypothesis is true, status distance should have negative effect on reciprocity, that is both β_2 and β_3 are negative and significant.

If both status asymmetry and homophily hypotheses are true, individuals are more likely to reciprocate nominations from people with higher status, the negative effect of status distance on reciprocity should be smaller for ego > alter dyads than for ego <= alter dyads; that is, $\beta_2 < \beta_3 < 0$.

4.5 Results

In the sample composed of 10,498 Add Health participants with valid in-school network measures, wave I in-home survey information and friend nominations, there are 2196 best same sex friend nominations based on wave I in-home survey. 46.9% of the nominations (1030 ties) were reciprocated, which is higher than meanwhile compatible to the

40% from another study based on Add Health wave I all friend nominations (Mouw & Entwisle, 2006).

Table 13 presents descriptive statistics for the 2196 dyads based on best same-sex friend nominations. It shows that friendships are more likely to be initiated by people with lower status toward people with higher status, with regard to global network centrality and influence, in-degree, out-degree, SES, age, grade, GPA and PTV score. 83% of best friend nominations are directed to someone from the same racial group, and on average, the dyads have 3 activities together on a weekly basis.

As shown in table 14, three models are estimated using random effect model. Estimates in the table are relative risk ratios, interpreted as multiplicative effects of the relative risk of friendship reciprocity. For example, a coefficient of 0.79 for centrality difference means that when the alter's centrality is greater than or equal to ego's centrality, the relative risk for alter to reciprocate ego's nomination is 0.79 if the centrality difference between alter and ego increases for one standard deviation.

Model 1 examines the effects of social network position and distance on friendship reciprocity without controls. The result doesn't suggest that alter tends to reciprocate nominations from ego with higher status, with regard to either centrality or influence. While status difference does matter, the likelihood of reciprocity decreases as the status distance increases, with regard to both centrality and influence, but only for nominations directed from ego to alter with equal or higher status.

Model 2 includes status asymmetry and status distance with regard to out-degree, in-degree, age, grade, SES, GPA, and PVT score as controls. The findings are largely consistent with model 1 except for that the effect of centrality difference for ego <= alter dyads turns

marginally significant and the effects of influence difference for ego > alter dyads becomes significant. The results based on controls tell the consistent story which suggests homophily instead of status asymmetry as the mechanism for friendship reciprocity, and the differences between ego and alter regarding in-degree, grade and SES exert negative effects on friendship reciprocity.

Model 3 incorporates more individual, relational and school level controls presented in the descriptive table. The results are consistent with model 2 after adding additional controls. In the three models, we observe a consistent pattern of homophily with respect to global network influence, net of the effects of control variables. Especially in model 3, the effects of influence difference are rather close for ego <= alter dyads and ego > alter dyads, which suggests homophily but not status asymmetry.

The study then uses mixed logit model to examine alter's choices of nominations to reciprocate. The specification takes a random coefficient form with the coefficient β varying over alters. The beta coefficients can logically take either sign, and each of the beta coefficients is given an independent normal distribution with mean and standard deviation that are estimated. Simulation was performed to yield estimates and the results are given in table 15. The standard deviations for few random coefficients are significant, suggesting that these coefficients do not vary much in the population. The findings from mixed logit model are consistent with those from random effect models. Status distances regarding to influence domain, in-degree and grade exhibit negative impacts on friendship reciprocity, but only for ego <= alter dyads, which supports the homophily hypothesis.

4.6 Conclusion

Friendship relations are often assumed to be mutual, but not all friend nominations

are reciprocated. The imbalanced nature of friendship is important to study as it addresses the crucial qualitative feature of the interpersonal relationship (Hartup, 1996). Friendship formation involves an initiation-response process, although friendships are often initiated by people with lower status toward people with higher status, which presents a pattern of status asymmetry, when it comes to friendship reciprocation, it may suggest a different story.

Both mixed model and discrete choice model are used to test the hypotheses derived from homophily and status asymmetry perspectives. All the results consistently direct to homophily as the major mechanism for friendship reciprocity. Different from friendship initiation which presents the pattern of both homophily and status asymmetry, the response process – friendship reciprocity – majorly presents the pattern of homophily. This might be explained by several reasons: it first involves the cost to maintain a friendship. Compared with friendship initiation, friendship reciprocity would directly lead to a mutual commitment to a relationship, and closeness in status is always associated with ease of maintaining the relationship, thus homophily is more serious a concern for friendship reciprocity than initiation. Other aspects involve the functions of friendship, people expect support from friends, mutual friendships, as the results of friendship reciprocity, are the intimate relationships that provide more support than unilateral friendships, thus status proximity, which suggests greater likelihood of providing support, is important when it comes to friendship reciprocity decision.

Table 1. Background Characteristics of 2094 Roommate Study Participants Who Were Paired with Roommates

Variable	Percentage/Mean
Gender	
Male	0.3792
Female	0.6208
Age	19.42
Race/Ethnicity	
Hispanic	0.074
White	0.6576
African American	0.1237
Asian	0.0688
American Indian	0.0019
Other	0.0124
Multiracial	0.0616
Mother's Education	
Less than high school	0.0168
High school graduate	0.2364
College graduate	0.3681
More than college	0.3787
Median Family Income	$100,000-150,000
Income Level	
Upper (>=500,000)	0.033
Upper Middle (100,000 - 500,000)	0.4892
Lower Middle (50,000 - 100,000)	0.3222
Working (< 50, 000)	0.1556
GPA in past Fall semester	3.23
Sample Size	2094

Table 2. Descriptive Statistics for Dependent Variables of 2094 Roommate Study Participants Who Were Paired with Roommates

Behaviors	Description or Coding Definition		College Behaviors (1)		
	(days per month)		Mean	Std.	N
Self-Pray	Prayed by yourself alone	Overall	7.15	8.37	2078
		Male	5.66	7.79	788
		Female	8.06	8.58	1290
Religious Service	Attended a religious service	Overall	1.6	2.97	2073
		Male	1.32	2.75	786
		Female	1.76	3.09	1287
Physical Exercise	Exercised or participated in physical activity for at least 20 minutes that made you sweat and breathe hard	Overall	9.16	6.86	2078
		Male	10.27	6.74	788
		Female	8.47	6.84	1290
Drinking	Drank alcoholic beverages	Overall	3.52	4.48	2083
		Male	4.5	5.06	790
		Female	2.92	3.97	1293
Binge Drinking	Drank five (four for female) or more drinks in a row	Overall	2.03	3.57	2084
		Male	2.72	4.14	787
		Female	1.61	3.1	1297
Smoking	Smoked cigarettes	Overall	0.98	3.73	2087
		Male	1.41	4.47	790
		Female	0.72	3.17	1297
Marijuana use	Used marijuana	Overall	0.89	3.27	2082
		Male	1.55	4.33	787
		Female	0.5	2.32	1295

(1)The first semester in college

Table 3. Effect of Roommates' High School Behavior (yes-or-no categorization) on Respondents' College Behavior

	Respondents' college behaviors (Days per month)						
Roommates' high school behaviors	Self-pray	Religious Service	Physical Exercise	Drinking	Binge Drinking	Smoking	Marijuana Use
Self-pray (Ref. = No)							
Yes	0.1178						
Religious Service (Ref. =No)							
Yes		-0.1323					
Physical Exercise (Ref. = No)							
Yes			0.2726				
Drinking (Ref. = No)							
Yes				0.3759*			
Binge Drinking (Ref. = No)							
Yes					0.3702**		
Smoking (Ref. = No)							
Yes						0.02263	
Marijuana Use (Ref. = No)							
Yes							-0.2745

Note: All regressions include controls for respondent's high school behavior, background characteristics of roommate and respondent (gender, race/ethnicity, GPA, family income and mother's education), and fixed effects including smoking status, residential type, campus location and two-way interaction of the three variables. *** p-value < 0.001; ** p-value <0.01; * p-value < 0.05.

Table 4. Effect of Roommates' High School Behavior (above-or-below median categorization) on Respondents' College Behavior

	Respondents' college behaviors (Days per month)						
Roommates' high school behaviors	Self-pray	Religious Service	Physical Exercise	Drinking	Binge Drinking	Smoking	Marijuana Use
Self-pray (Ref. = Below median)							
Above median	0.1582						
Religious Service (Ref. = Below median)							
Above median		-0.07565					
Physical Exercise (Ref. = Below median)							
Above median			0.7786**				
Drinking (Ref. = Below median)							
Above median				0.3759*			
Binge Drinking (Ref. = Below median)							
Above median					0.3702**		
Smoking (Ref. = Below median)							
Above median						0.02263	
Marijuana Use (Ref. = Below median)							
Above median							-0.2745

Note: All regressions include controls for respondent's high school behavior, background characteristics of roommate and respondent (gender, race/ethnicity, GPA, family income and mother's education), and fixed effects including smoking status, residential type, campus location and two-way interaction of the three variables. *** p-value < 0.001; ** p-value <0.01; * p-value < 0.05.

Table 5. Effect of Roommates' High School Behavior (three categorizations) on Respondents' College Behavior

	Respondents' college behaviors (Days per month)						
Roommates' high school behaviors	Self-pray	Religious Service	Physical Exercise	Drinking	Binge Drinking	Smoking	Marijuana Use
Self-pray (Ref. = none)							
Occasional	-0.325						
Frequent	0.1937						
Religious Service (Ref. =none)							
Occasional		-0.1513					
Frequent		0.4231*					
Physical Exercise (Ref. = none)							
Occasional			-0.7067				
Frequent			0.8725**				
Drinking (Ref. = none)							
Occasional				0.7247***			
Frequent				-0.1193			
Binge Drinking (Ref. = none)							
Occasional					0.3251*		
Frequent					-0.9363		
Smoking (Ref. = none)							
Occasional						0.223	
Frequent						0.078	
Marijuana Use (Ref. = none)							
Occasional							-0.239
Frequent							-0.472

Note: All regressions include controls for respondent's high school behavior, background characteristics of roommate and respondent (gender, race/ethnicity, GPA, family income and mother's education), and fixed effects including smoking status, residential type, campus location and two-way interaction of the three variables. *** p-value < 0.001; ** p-value <0.01; * p-value < 0.05.

Table 6. Effect of Roommates' High School Behavior (three categorizations) on Respondents' College Behavior, for Males

	Respondents' college behaviors (Days per month)						
Roommates' high school behaviors	Self-pray	Religious Service	Physical Exercise	Drinking	Binge Drinking	Smoking	Marijuana Use
Self-pray (Ref. = none)							
Occasional	-0.262						
Frequent	-0.213						
Religious Service (Ref. =none)							
Occasional		-0.014					
Frequent		0.136					
Physical Exercise (Ref. = none)							
Occasional			-3.578				
Frequent			**0.982***				
Drinking (Ref. = none)							
Occasional				**0.695***			
Frequent				-0.878			
Binge Drinking (Ref. = none)							
Occasional					0.406		
Frequent					-0.547		
Smoking (Ref. = none)							
Occasional						0.318	
Frequent						-0.166	
Marijuana Use (Ref. = none)							
Occasional							-0.253
Frequent							-0.649

Note: All regressions include controls for respondent's high school behavior, background characteristics of roommate and respondent (gender, race/ethnicity, GPA, family income and mother's education), and fixed effects including smoking status, residential type, campus location and two-way interaction of the three variables. *** p-value < 0.001; ** p-value <0.01; * p-value < 0.05.

Table 7. Effect of Roommates' High School Behavior (three categorizations) on Respondents' College Behavior, for Females

	Respondents' college behaviors (Days per month)						
Roommates' high school behaviors	Self-pray	Religious Service	Physical Exercise	Drinking	Binge Drinking	Smoking	Marijuana Use
Self-pray (Ref. = none)							
Occasional	-0.3658						
Frequent	0.4118						
Religious Service (Ref. =none)							
Occasional		-0.2782					
Frequent		**0.5851***					
Physical Exercise (Ref. = none)							
Occasional			0.2534				
Frequent			0.6421				
Drinking (Ref. = none)							
Occasional				**0.3887***			
Frequent				-0.01087			
Binge Drinking (Ref. = none)							
Occasional					0.2703		
Frequent					-0.9036		
Smoking (Ref. = none)							
Occasional						0.005195	
Frequent						-0.1844	
Marijuana Use (Ref. = none)							
Occasional							-0.2912
Frequent							-0.2068

Note: All regressions include controls for respondent's high school behavior, background characteristics of roommate and respondent (gender, race/ethnicity, GPA, family income and mother's education), and fixed effects including smoking status, residential type, campus location and two-way interaction of the three variables. *** p-value < 0.001; ** p-value <0.01; * p-value < 0.05.

Table 8. Behavioral Correlations between Roommates Over Time

	All sample (N=2094)		Students in year 2 and above and no longer living with roommate (N=1176)			Students in year 2 and above and still living with roommate (N= 278)		
Behaviors	High school	1st fall semester	High school	1st fall semester	last fall semester	High school	1st fall semester	last fall semester
Drinking	0.000851	0.009125***	0.005923	0.04901*	0.001257	-0.00212	0.02786***	0.01943***
Binge Drinking	0.001006	0.007951**	0.0069	0.01074**	0.005633	0.00894	0.04262***	0.02625*
Smoking	-0.00275	0.000704	-0.00223	-0.00375	-0.00462	-0.01362	0.01649*	0.008981
Marijuana	-0.00643	-0.00775	-0.0043	-0.00754	-0.00102	-0.04811	-0.00578	0.008826

Note: All regressions include controls for prior behaviors of roommate and respondent, background characteristics of roommate and respondent (gender, race/ethnicity, GPA, family income and mother's education), and fixed effects including smoking status, residential type, campus location and two-way interaction of the three variables. *** p-value < 0.001; ** p-value <0.01; * p-value < 0.05

Table 9. Descriptive Statistics for AddHealth and Roommate Study Analytical Samples

	AddHealth		**Roommate Study**
	DNA Sample	Full Sample (wave III)	
Age at wave I			
16 or older	52.18%	53.18%	100%
less than 16	47.82%	46.82%	0
Gender			
Male	48.61%	47.23%	36.47%
Female	51.39%	52.77%	63.53%
Race/Ethnicity			
Hispanic	14.49%	16.18%	8.13%
White	57.40%	51.83%	66.57%
African American	16.74%	20.34%	10.27%
Asian	6.87%	6.58%	7.15%
Native American	0.18%	0.55%	0.19%
Other	0.88%	0.78%	1.24%
Multiracial	3.44%	3.74%	6.45%
Mother's education			
Less than high school	16.94%	17.33%	1.37%
High school graduate or GED	39.57%	38.67%	24.86%
College	35.66%	34.49%	36.15%
More than college	7.83%	9.51%	37.62%
Sample size	2281	15197	2080

Table 10. Number of Friend Pairs from Addhealth and Roommate Study Analytical Samples

	Addhealth	Roommate Study
Reciprocal friends		
In-School	123	
Wave I	81	-
Wave II	81	-
Overall (no overlap)	206	212
Non-Reciprocal friends		
In-School	363	-
Wave I	266	-
Wave II	212	-
Overall (no overlap)	689	214
All friends	895	426

Table 11. Average Genetic Correlations for Friend Pairs, Simulated Distributions of Average Genetic Correlations for Random Pairs and Significance Tests

			random pairs (n of replications =3000)		
	n of pairs	average rho value	mean	s.d	P value
Addhealth					
Reciprocal friends	206	0.047	0.0023	0.0074	1.12E-11
non-reciprocal friends	689	0.039	0.0027	0.0055	1.02E-06
Roommate Study					
Reciprocal friends	212	0.0075	0.0034	0.0069	0.26
non-reciprocal friends	214	0.0048	0.0032	0.0064	0.39
Quality control: Addhealth sample of known relatedness					
full siblings	556	0.52	0.0027	0.0059	0
MZ twins	19	0.99	0.0014	0.0090	0
DZ twins	186	0.52	0.0021	0.0079	0

Table 12. Average Genetic Correlations for Reciprocal and Non-reciprocal Friends in AddHealth: All, In-school, Wave I and Wave II.

	Reciprocal Friends			Non-Reciprocal Friends			Difference
	n of pairs	average rho	S.D.	n of pairs	average rho	S.D.	P value (two sample T test)
All	206	0.047	0.182	689	0.039	0.179	0.069
In-school	123	0.08	0.188	363	0.035	0.175	0.035
Wave I	81	0.062	0.206	266	0.044	0.178	0.095
Wave II	81	0.05	0.191	212	0.042	0.177	0.386

Table 13. Descriptive Statistics for 2196 Dyads Based on Best Same-sex Friend Nominations

	Mean/Percentage
Dependent Variable	
Reciprocity (1)	46.90%
Independent Variables	
Ego centrality > Alter centrality (2)	47.13%
Centrality difference	
Ego <= Alter	0.79
Ego >Alter	0.77
Ego influence > Alter influence	9.65%
Influence difference	
Ego <= Alter	0.16
Ego >Alter	0.52
Controls	
Ego out-degree>Alter out-degree	40.94%
Out-degree difference	
Ego <= Alter	2.39
Ego >Alter	3.13
Ego in-degree>Alter in-degree	35.25%
In-degree difference	
Ego<=Alter	3.32
Ego >Alter	3.28
Ego age > Alter age	23.18%
Age difference	
Ego <= Alter	0.36
Ego >Alter	1.16
Ego grade > Alter grade	10.43%
Grade difference	
Ego <= Alter	0.16
Ego >Alter	1.18
Ego SES > Alter SES	33.42%
SES difference	
Ego <= Alter	0.74
Ego >Alter	1.48
Ego GPA > Alter GPA	42.58%
GPA difference	
Ego <= Alter	0.47
Ego >Alter	0.57
Ego PVT score > Alter PVT score	48.77%
PVT score difference	
Ego <= Alter	11.37
Ego >Alter	11.81

Same race	82.74%
Ego's centrality	0.15
Ego's influence	0.15
Ego's out-degree	5.15
Ego's in-degree	5.27
Ego's gender (1=Male)	44.99%
Ego's age	16.11
Ego's grade	9.73
Ego's race	
Asian	8.57%
Black	15.45%
Hispanic(3)	15.18%
Multiracial	3.46%
Native	0.46%
Other	0.73%
White	56.15%
Ego's SES	
Less than High School	9.46%
High School Graduate or GED	29.73%
Some College	22.41%
College Graduate	25.40%
More than College	13%
Ego's GPA	3.19
Ego's PVT score	101.82
Ego's local network density	0.31
School characteristics	
Index of mutuality in school network	0.39
School size	
125 or fewer students	2.87%
126-350 students	12.72%
351-775 students	22.30%
more than 775 students	62.11%
Metropolitan location	
Urban	18.06%
Suburban	49.25%
Rural	32.69%
Friendship characteristics	
Number of activities with best same-sex friend	3.29
N	2196

(1) Friendship nominations are based upon wave I in-home survey
(2) All global and local network measures are based on in-school survey
(3) all other racial categories are non-Hispanic

Table 14. Effects of Status Asymmetry and Distance on Friendship Reciprocity: Random Effect Models

	Model 1	Model 2	Model 3
Independent Variables			
Ego centrality > Alter centrality	0.95	0.94	0.86
Centrality difference			
Ego <= Alter	0.79**	0.73+	0.76
Ego >Alter	0.92	1.03	1.05
Ego influence > Alter influence	0.88	0.83	1.08
Influence difference			
Ego <= Alter	0.46***	0.55***	0.69*
Ego >Alter	0.75	0.72*	0.70*
Controls			
Ego out-degree>Alter out-degree		1.11	1.14
Out-degree difference			
Ego <= Alter		1.06	1.07
Ego >Alter		0.94	0.94
Ego in-degree>Alter in-degree		1.07	1.09
In-degree difference			
Ego<=Alter		0.89***	0.88***
Ego >Alter		1.04	1.05
Ego age > Alter age		0.93	0.87
Age difference			
Ego <= Alter		1.19	1.14
Ego >Alter		1.21	1.20
Ego grade > Alter grade		1.26	1.46
Grade difference			
Ego <= Alter		0.59***	0.56***
Ego >Alter		0.68	0.58
Ego SES > Alter SES		1.19	1.19
SES difference			
Ego <= Alter		0.87*	0.85
Ego >Alter		1.14	1.24
Ego GPA > Alter GPA		1.24	1.17
GPA difference			
Ego <= Alter		0.80	0.94
Ego >Alter		0.78	0.81
Ego PVT score > Alter PVT score		1.01	1.04
PVT score difference			
Ego <= Alter		1.002	1.01

Ego >Alter		1.0009	0.995
N	2196	1891	1859
-2LL	2943.93	2433.72	2326.68

Note: Estimates are presented in exponential form.
For model 3, the results for controls including same race, ego's centrality, influence, out-degree, in-degree, gender, age, grade, race, SES, GPA, PVT score, ego's local network density, index of school network mutuality, school size, metropolitan location and number of activities with best same-sex friend are not presented.

Table 15. Effects of Status Asymmetry and Distance on Friendship Reciprocity: Discrete Choice Model

Variable	Mean of beta		S.D. of beta	
Ego centrality > Alter centrality	-0.065		0.18	
Centrality difference				
Ego <= Alter	-0.34		0.16	
Ego >Alter	0.027		0.21	***
Ego influence > Alter influence	-0.16		0.19	
Influence difference				
Ego <= Alter	-0.63	***	0.13	
Ego >Alter	-0.13		0.17	
Ego outdegree>Alter outdegree	0.11		0.2	
Outdegree difference				
Ego <= Alter	0.057		0.059	
Ego >Alter	-0.062		0.071	
Ego indegree>Alter indegree	0.077		0.15	
Indegree difference				
Ego<=Alter	-0.13	***	0.02	
Ego >Alter	-0.044		0.031	
Ego age > Alter age	-0.076		0.13	
Age difference				
Ego <= Alter	0.17		0.11	
Ego >Alter	0.28		0.25	
Ego grade > Alter grade	0.27		0.43	***
Grade difference				
Ego <= Alter	-0.54	*	0.15	
Ego >Alter	-0.28		0.21	
Ego SES > Alter SES	-0.27		0.19	
SES difference				
Ego <= Alter	-0.16		0.19	
Ego >Alter	0.13		0.22	*
Ego GPA > Alter GPA	0.29		0.21	
GPA difference				
Ego <= Alter	-0.23		0.14	
Ego >Alter	-0.25		0.17	
Ego PVT score > Alter PVT score	0.02		0.14	
PVT score difference				
Ego <= Alter	0.0012		0.007	
Ego >Alter	0.0011		0.009	
Log Likelihood	-1266.79			

Figure 1. Changes in Behavioral Patterns from High School to First Semester in College of 2094 Roommate Study Participants Who Were Paired with Roommates

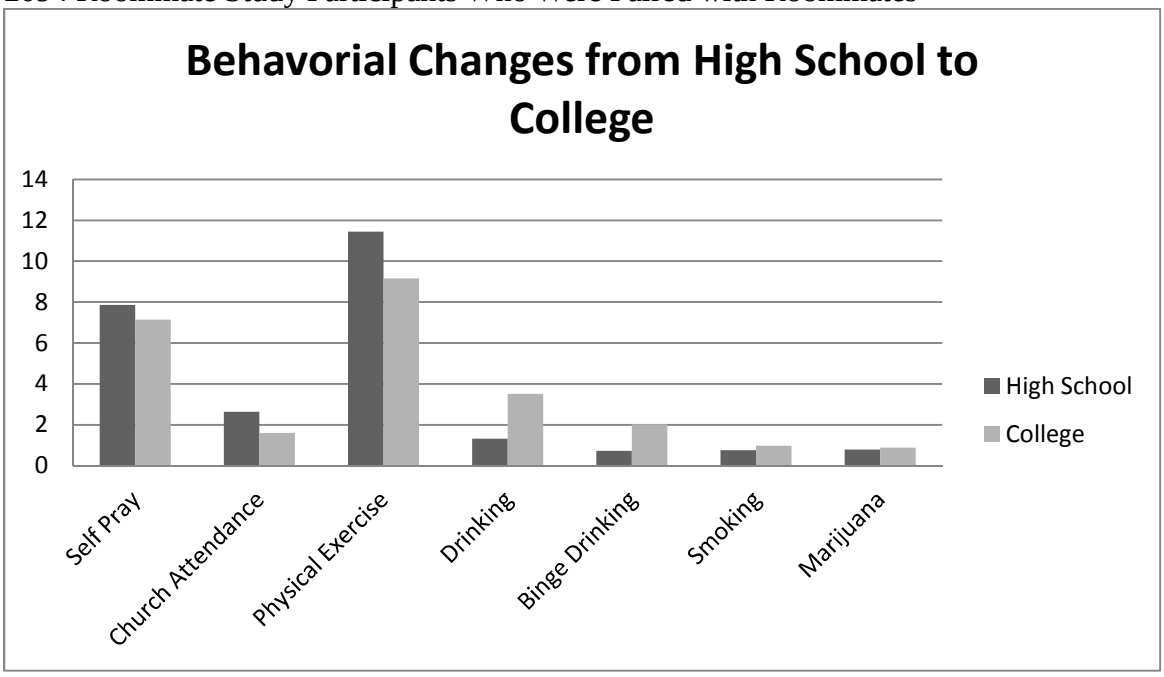


Figure 2. Three Levels of Mechanisms from Behavior Related Genes to Genetic Homophily among Friends

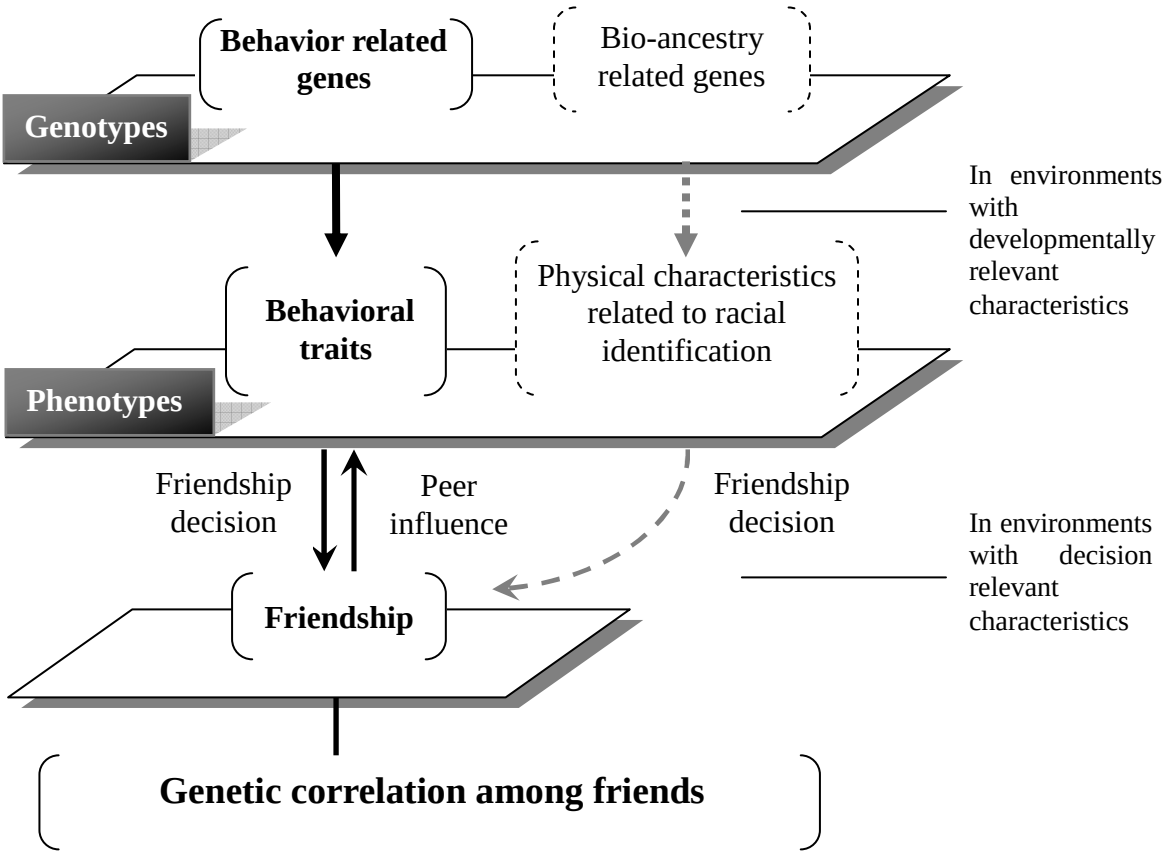
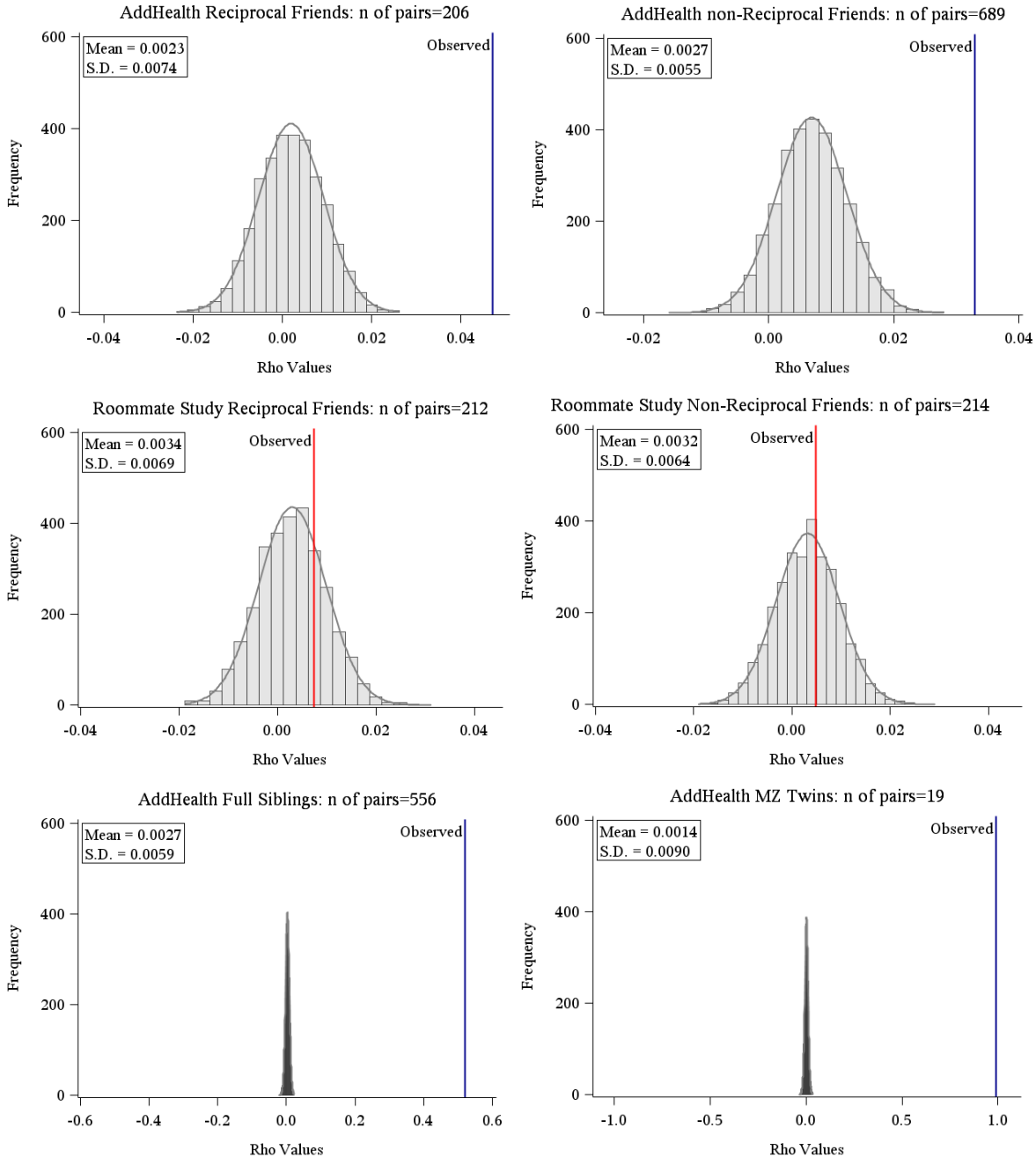


Figure 3. Simulation Tests of Genetic Association in Friends, Full Siblings, MZ and DZ Twins



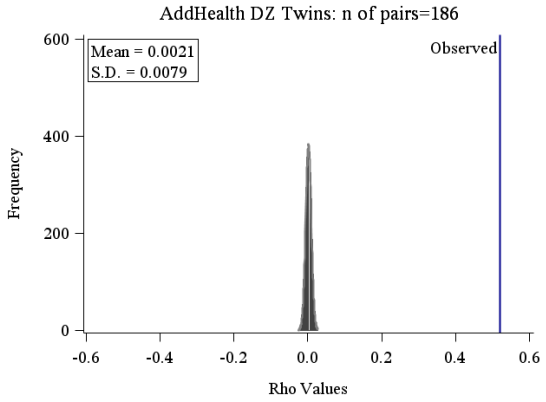
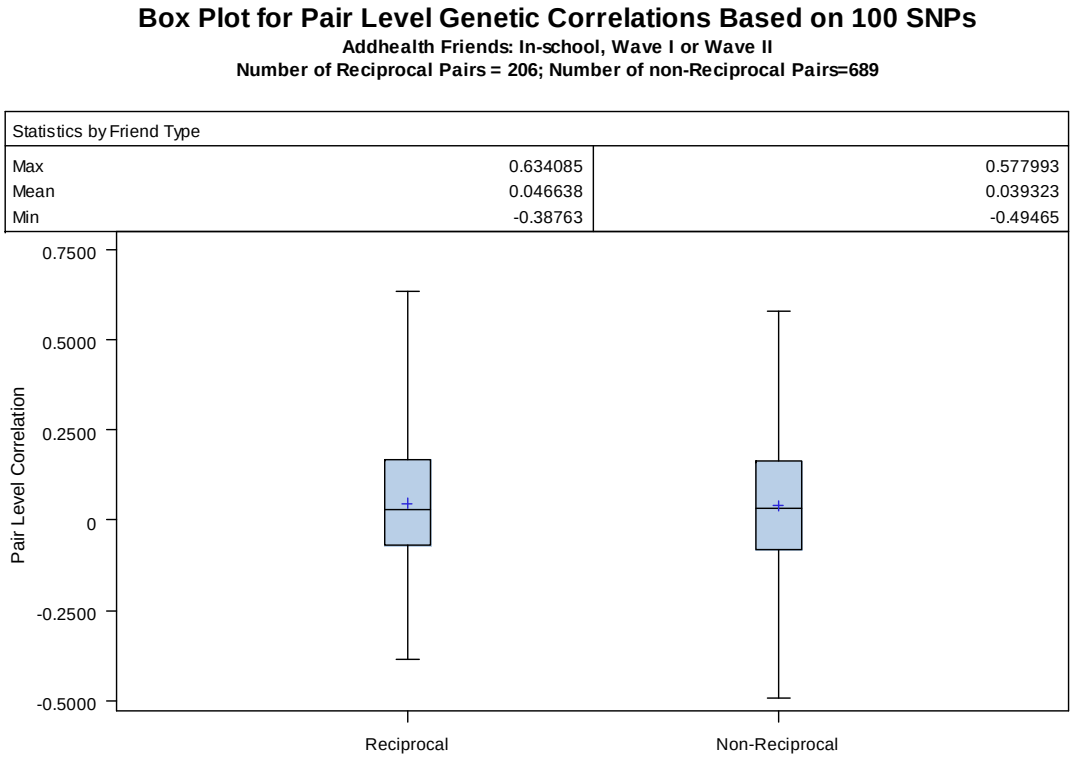

AddHealth DZ Twins: n of pairs=186
Mean = 0.0021
S.D. = 0.0079
Observed
Frequency
600
400
200
0
-0.6
-0.4
-0.2
0.0
0.2
0.4
0.6
Rho Values

Figure 4. Box Plot for Rho Values for AddHealth Reciprocal and Non-reciprocal Friends: All Waves Combined



REFERENCES

Akers, R. L., Krohn, M. D., Lanza-Kaduce, L., & Radosevich, M. (1979). Social learning and deviant behavior: a specific test of a general theory. *American Sociological Review*, *44*(4), 636–655.

Allen, G., Eriksson, A. W., Fellman, J., Parisi, P., & Vandenberg, S. G. (1992). Twinning and the r/K reproductive strategy: a critique of Rushton's theory. *Acta Geneticae Medicae Et Gemellologiae*, *41*(1), 73–83.

Aukett, R., Ritchie, J., & Mill, K. (1988). Gender differences in friendship patterns. *Sex Roles*, *19*(1-2), 57–66.

Bavelas, A. (1948). A mathematical model for group structures. *Human Organization*, (7), 16–30.

Belle, D. (1989). Gender differences in children's social networks and supports. *Children's Social Networks and Social Supports*. New York: Wiley.

Biesanz, J. C., West, S. G., & Millevoi, A. (2007). What do you learn about someone over time? The relationship between length of acquaintance and consensus and self-other agreement in judgments of personality. *Journal of personality and social psychology*, *92*(1), 119–135.

Billy, J. O. G., & Udry, J. R. (1985). Patterns of adolescent friendship and effects on sexual behavior. *Social Psychology Quarterly*, 27–41.

Blackwell, D. L., & Lichter, D. T. (2004). Homogamy Among Dating, Cohabiting, and Married Couples. *The Sociological Quarterly*, *45*(4), 719–737.

Blau, P. M. (1974). Presidential Address: Parameters of Social Structure. *American Sociological Review*, *39*(5), 615–635.

Boardman, J. D., Domingue, B. W., & Fletcher, J. M. (2012). How social and genetic factors predict friendship networks. *Proceedings of the National Academy of Sciences*, *109*(43), 17377–81.

Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social Networks*, *29*(4), 555–564.

Bonacich, Phillip. (1987). Power and Centrality□: A Family of Measures. *The American Journal of Sociology*, *92*(5), 1170–1182.

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

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

Brechwald, W. a., & Prinstein, M. J. (2011). Beyond Homophily: A Decade of Advances in Understanding Peer Influence Processes. *Journal of Research on Adolescence*, *21*(1), 166–179.

Brown, B. B., Lohr, M. J., & Trujillo, C. M. (1990). Multiple Crowds and Multiple Lifestyles: Adolescents' Perceptions of Peer Group Characteristics. In E. R. Muuss (Ed.), *Adolescent Behavior and Society: A Book of Readings* (pp. 30–36). New York, NY: Random House.

Bryk, A. S. (1996). Lessons from catholic high schools on renewing our educational institutions. In T. H. McLaughlin, J. O'Keefe, & B. O'Keefe (Eds.), *The Contemporary Catholic School: Context, Identity, and Diversity.* London: Falmer Press.

Burgess, E. W., & Wallin, P. (1943). Homogamy in Social Characteristics. *American Journal of Sociology*, *49*(2), 109–124.

Byrne, D. (1971). *The Attraction Paradigm*. New York: Academic Press.

Cain, D. P., & Vanderwolf, C. H. (1990). A critique of Rushton on race, brain size and intelligence. *Personality and Individual Differences*, *11*(8), 777–784.

Casciaro, T., & Lobo, M. S. (2008). When Competence Is Irrelevant: The Role of Interpersonal Affect in Task-Related Ties. *Administrative Science Quarterly*, *53*(4), 655–684.

Christakis, N. A., & Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *The New England journal of medicine*, *357*(4), 370–379.

Christakis, N. A., & Fowler, J. H. (2008). The collective dynamics of smoking in a large social network. *The New England journal of medicine*, *358*(21), 2249–2258.

Christakis, N. A., & Owler, J. H. (2012). Social contagion theory: examining dynamic social networks and human behavior. *Statistics in medicine*, (November 2011).

Cole, T., & Teboul, J. B. (2004). Non-Zero-Sum Collaboration, Reciprocity, and the Preference for Similarity: Developing and Adaptive Model of Close Relational Functioning. *Personal Relationships*, *11*, 135–60.

Crosnoe, R., Frank, K., & Mueller, A. S. (2008). Gender, Body Size and Social Relations in American High Schools. *Social Forces*, *86*(3), 1189–1216.

Daniels, D., & Plomin, R. (1985). Differential experience of siblings in the same family. *Developmental Psychology*, *21*(5), 747–760.

Duncan, G., & Brooks-Gunn, J. (1997). *Consequences of growing up poor*. New York, NY: Russell Sage Foundation.

Duncan, G. J., Boisjoly, J., Kremer, M., Levy, D. M., & Eccles, J. (2005). Peer Effects in Drug Use and Sex Among College Students. *Journal of Abnormal Child Psychology*, *33*(3), 375–385.

Dunn, L. M. (1981). *The Peabody Picture Vocabulary Test-Revised*. Circle Pines, MN: American Guidance Service.

Ellison, C. G., & Levin, J. S. (1998). The Religion-Health Connection: Evidence, Theory, and Future Directions. *Health Education & Behavior*, *25*(6), 700–720.

Emerson, R. M. (1976). Social Exchange Theory. (J. Delamater, Ed.)*Annual Review of Sociology*, *2*(1976), 335–362.

Fine, G. A. (1980). The Natural History of Preadolescent Male Friendship Groups. In H. C. Foot, A. J. Chapman, & J. R. Smith. (Eds.), *Friendship and Social Relations in Children* (pp. 293–32). New York, NY: Wiley.

Fletcher, J. M., & Frisvold, D. E. (2011). College selectivity and young adult health behaviors. *Economics of Education Review*, *30*(5), 826–837.

Foster, G. (2006). It's not your peers, and it's not your friends: Some progress toward understanding the educational peer effect mechanism. *Journal of Public Economics*, *90*(8-9), 1455–1475.

Fowler, J. H., Settle, J. E., & Christakis, N. A. (2010). Correlated genotypes in friendship networks. *Proceedings of the National Academy of Sciences*, (24), 1–5.

Freeman, L. C. (1992). Filling in the Blanks: A Theory of Cognitive Categories and the Structure of Social Affiliation. *Social Psychology Quarterly*, *55*(2), 118–127.

Frudakis, T., Venkateswarlu, K., Thomas, M. J., Gaskin, Z., Ginjupalli, S., Gunturi, S., Ponnuswamy, V., et al. (2003). A classifier for the SNP-based inference of ancestry. *Journal of Forensic Sciences*, *48*(4), 771–782.

Gould, R. V. (2002). The Origins of Status Hierarchies: A Formal Theory and Empirical Test. *American Journal of Sociology*, *107*(5), 1143–1178.

Guo, G. (2006). Genetic similarity shared by best friends among adolescents. *Twin research and human genetics : the official journal of the International Society for Twin Studies*, *9*(1), 113–21.

Guo, G., Li, Y., Owen, C., Duncan, G., Cai, T., Fu, Y., & Wang, H. (2012). Peer Influences, Self Prior Behavior History and Youth Alcohol Use: A Natural Experiment on a College Campus and a Survey Study. *Submitted to Social Science and Research*.

Haas, S. a, Schaefer, D. R., & Kornienko, O. (2010). Health and the structure of adolescent social networks. *Journal of health and social behavior*, *51*(4), 424–39.

Hallinan, M. T. (1978). The Process of Friendship Formation. *Social Networks*, *1*(2), 193–210.

Hamilton, W. D. (1964). The Genetical Evolution of Social Behaviour I and II. *Journal of Theoretical Biology*.

Hartup, W. W. (1996). Presidential Address The Company They Keep : Friendships Their Developmental Significance. *Significance*, *67*(1), 1–13.

Haynie, D. L., & Osgood, D. W. (2005). Reconsidering Peers and Delinquency: How do Peers Matter? *Social Forces*, *84*(2), 1109–1130.

Haynie, Dana L. (2001). Delinquent Peers Revisited : Does Network Structure Matter ? *American Journal of Sociology*, *106*(4).

Heller, P. E., & Wood, B. (1998). The process of intimacy: similarity, understanding and gender. *Journal of marital and family therapy*, *24*(3), 273–288.

Jacobson, K. C., & Rowe, D. C. (1998). Genetic and shared environmental influences on adolescent BMI: interactions with race and sex. *Behavior Genetics*, *28*(4), 265–278.

Johnson, P. B. (1989). Reactions, expectancies, and college students' drinking. *Psychological Reports*, *65*(3 Pt 2), 1245–1246.

Johnson, R. (1979). *Juvenile delinquency and its origins*. Cambridge, UK: Cambridge University Press.

Johnston, L. D., O'Malley, P. M., Bachman, J. G., & Schulenberg, J. E. (2008). Monitoring the Future national survey results on drug use, 1975-2007. Volume II: College students and adults ages 19-45. *Social Research*. National Institute on Drug Abuse.

Kandel, D. B. (1985). On Processes of Peer Influences in Adolescent Drug Use. *Advances in Alcohol and Substance Abuse*, (4), 139–162.

Katz, L., & Powell, J. H. (1955). Measurement of the Tendency Toward Reciprocation of Choice. *Sociometry*.

Klein, H. (1992). College student's attitudes toward the use of alcoholic beverages. *Journal of Alcohol and Drug Education*, *37*, 35–52.

Kremer, M., & Levy, D. (2008). Peer Effects and Alcohol Use among College Students. *The Journal of Economic Perspectives*, *22*(3), 189–206.

Lareau, A. (2000). *Home Advantage: Social Class and Parental Intervention in Elementary Education.* Lanham: Rowmand and Littlefield.

Larson, R. W., & Verma, S. (1999). How children and adolescents spend time across the world: work, play, and developmental opportunities. *Psychological Bulletin*, *125*(6), 701–736.

Laursen, B. (Ed.). (1993). Close friendships in adolescence. *New Directions for Child Development*. San Francisco: Jossey-Bass.

Lazersfeld, P. F., & Merton, R. K. (1954). Friendship as Social Process: A Substantive and Methodological Analysis. In M. Berger (Ed.), *Freedom and Control in Modern Society* (pp. 18–66). New York: Van Nostrand.

Ledbetter, A. M., Griffin, E., & Sparks, G. G. (2007). Forecasting "friends forever": A longitudinal investigation of sustained closeness between best friends. *Personal Relationships*, *14*(2), 343–350.

Lewis, R. (2003). *Human Genetics: Concepts and Applications* (5th ed.). McGraw-Hill Higher Education.

Manski, C. F. (1993). Identification of Endogenous Social Effects□: The Reflection Problem. *The Review of Economic Studies*, *60*(3), 531–542.

McFadden, D., & Train, K. (2000). Mixed MNL models for discrete response. *Journal of Applied Econometrics*, *15*(5), 447–470.

McPherson, J. M., & Ranger-Moore, J. R. (1991). Evolution on a Dancing Landscape: Organizations and Networks in Dynamic Blau Space. *Social Forces*, *70*(1), 19–42.

Mcpherson, M., Smith-lovin, L., & Cook, J. M. (2001). Birds of A Feather□: Homophily in Social Networks. *Annual Review of Sociology*, *27*, 415–444.

Mead, G. H. (1934). *Mind, self, and society: from the standpoint of a social behaviourist*. (C. W. Morris, Ed.)*Chicago University of Chicago Press*. Chicago University Press.

Mears, D. P., Ploeger, M., & Warr, M. (1998). EXPLAINING THE GENDER GAP IN DELINQUENCY□: PEER INFLUENCE AND MORAL EVALUATIONS OF BEHAVIOR. *Journal Of Research In Crime And Delinquency*, *35*(3), 251–266.

Mendelson, M. J., & Kay, A. C. (2003). Positive Feelings in Friendship: Does Imbalance in the Relationship Matter? *Journal of Social and Personal Relationships*, *20*(1), 101–116.

Milner Jr., M. (2004). *Freaks, Geeks, and Cool Kids: American Teenagers, Schools, and the Culture of Consumption*. New York: Routledge.

Mollica, K. A., Gray, B., & Trevino, L. K. (2003). Racial Homophily and Its Persistence in Newcomers' Social Networks. *Organization Science*, *14*(2), 123–136.

Mouw, T. (2003). Social Capital and Finding a Job: Do Contacts Matter? *American Sociological Review*, *68*(6), 868–898.

Mouw, T. (2006). Estimating the Causal Effect of Social Capital: A Review of Recent Research. *Annual Review of Sociology*, *32*(1), 79–102.

Mouw, T., & Entwisle, B. (2006). Residential Segregation and Interracial Friendship in Schools. *American Journal of Sociology*, *112*(2), 394–441.

Newcomb, T. M. (1961). *The Acquaintance Process*. *The Acquaintance Process.* New York: Holt, Rinehart and Winston.

Olk, P. M., & Gibbons, D. E. (2010). Dynamics of Friendship Reciprocity Among Professional Adults. *Journal of Applied Social Psychology*, *40*(5), 1146–1171.

Owen, C. (2012). Peer effects as mediated by roommate relationship quality and friendship status. *Working Paper*.

Perry-Smith, J.□;, & Shalley, C. (2003). The social side of creativity: A static and dynamic social network perspective. *Academy of Management Review*, (28), 89–106.

Plickert, G., Côté, R. R., & Wellman, B. (2007). It's not who you know, it's how you know them: Who exchanges what with whom? *Social Networks*, *29*(3), 405–429.

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, *38*(8), 904–909.

Qian, Z., & Lichter, D. T. (2007). Social Boundaries and Marital Assimilation: Interpreting Trends in Racial and Ethnic Intermarriage. *American Sociological Review*, *72*(1), 68–94.

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

Rogers, E. M., & Bhowmik, D. K. (1970). Homophily-Heterophily: Relational Concepts for Communication Research. (L. L. Barker & R. J. Kibler, Eds.)*Public Opinion Quarterly*, *34*(4), 523–538.

Rosenberg, N. A., Pritchard, J. K., Weber, J. L., Cann, H. M., Kidd, K. K., Zhivotovsky, L. A., & Feldman, M. W. (2002). Genetic Structure of Human Populations. *Science*, *298*.

Roth, A. E. ., & Sotomayor, M. A. O. (1990). *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Cambridge, UK: Cambridge University Press.

Ruef, M., Aldrich, H. E., & Carter, N. M. (2003). The Structure of Founding Teams: Homophily, Strong Ties, and Isolation among U.S. Entrepreneurs. *American Sociological Review*, *68*(2), 195–222.

Rushton, J P, Russell, R. J. H., & Wells, P. A. (1985). Personality and genetic similarity theory. *Journal of Social Biological Structures*. Academic Press.

Rushton, J Philippe. (1989). Genetic similarity, human altruism, and group selection. *Behavioral and Brain Sciences*, *12*(03), 503.

Rushton, J Philippe. (2005). Ethnic nationalism, evolutionary psychology and Genetic Similarity Theory. *Nations and Nationalism*, *11*(4), 489–507.

Sacerdote, B. (2001). Peer Effects with Random Assignment□: Results for Dartmouth Roommates. *The Quarterly Journal of Economics*, *116*(2), 681–704.

Schaefer, D. R. (2012). Homophily Through Nonreciprocity: Results of an Experiment. *Social Forces*, *90*(4), 1271–1295.

Sijtsema, J. J., Ojanen, T., Veenstra, R., Lindenberg, S., Hawley, P. H., & Little, T. D. (2009). Forms and Functions of Aggression in Adolescent Friendship Selection and Influence: A Longitudinal Social Network Analysis. *Social Development*, *19*(3), 515–534.

Simpson, D. B., Cloud, D. S., Newman, J. L., & Fuqua, D. R. (2008). Sex and Gender Differences in Religiousness and Spirituality. *Psychology*, *36*(1), 42–52.

Smith, K. P., & Christakis, N. a. (2008). Social Networks and Health. *Annual Review of Sociology*, *34*(1), 405–429.

Stinebrickner, R., & Stinebrickner, T. (2006). What can be learned about peer effects using college roommates? Evidence from new survey data and students from disadvantaged backgrounds. *Journal of Public Economics*, *90*(8-9), 1435–1454.

Sutherland, Edwin Hardin; Cressey, D. R. (1955). *Principles of Criminology* (5th ed.). J.B. Lippincott.

Train, K. E. (2002). *Discrete Choice Methods with Simulation*. Cambridge, UK: Cambridge University Press.

Turkheimer, E. (2000). Three Laws of Behavior Genetics and What They Mean CAUSATION IN. *October*, *9*(5), 160–164.

Vaquera, E., & Kao, G. (2008). Do You Like Me as Much as I Like You? Friendship Reciprocity and Its Effects on School Outcomes among Adolescents. *Social science research*, *37*(1), 55–72.

Walster, E., Aronson, V., Abrahams, D., & Rottman, L. (1966). Importance of physical attractiveness in dating behavior. *Journal of Personality and Social Psychology*, *4*(5), 508–516.

Wasserman, S. ., & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge University Press.

Yakusheva, O., Kapinos, K., & Weiss, M. (2011). Peer effects and the freshman 15: evidence from a natural experiment. *Economics and human biology*, *9*(2), 119–32. doi:10.1016/j.ehb.2010.12.002

Zeng, Z., & Xie, Y. (2008). A preference-opportunity-choice framework with applications to intergroup friendship. *AJS; American journal of sociology*, *114*(3), 615–48.

Zimmerman, D. J. (2003). Peer Effects in Academic Outcomes: Evidence from A Natural Experiment. *The Review of Economics and Statistics*, *85*(February), 9–23.