# From Ties to Gains?
# Evidence on Connectedness, Skill Acquisition, and Diversity*

Philip Babcock
University of California, Santa Barbara

March 1, 2006

ABSTRACT

This paper uses micro-level data on social networks in middle and secondary schools to estimate effects of connectedness on education attainment outcomes and the association between racial diversity and connectedness. The analysis addresses concerns about unobserved neighborhood and school-level heterogeneity by using within-school variation between grade cohorts to identify effects of connectedness. Results include 2 main findings: 1) Being part of a more connected cohort within a given secondary or middle school is associated with significantly higher years of schooling attained and higher probability of having attended college, 7 years later; 2) Being part of a more racially diverse grade cohort, within a given school, is associated with significantly lower levels of connectedness—rare micro-level evidence to augment existing cross-region evidence on ethnic fractionalization and disconnectedness.

*The author thanks Julian Betts, Vince Crawford, Roger Gordon, and Gordon Hanson for advice and assistance. This work also benefited from my conversations with Kelly Bedard, Hoyt Bleakley, Julie Cullen, Mindy Marks, Valerie Ramey, James Rauch, and Joel Sobel. This research uses data from Add Health, a program project designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris, and funded by a grant P01-HD31921 from the National Institute of Child Health and Human Development, with cooperative funding from 17 other agencies. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Persons interested in obtaining data files from Add Health should contact Add Health, Carolina Population Center, 123 W. Franklin Street, Chapel Hill, NC 27516-2524 (addhealth@unc.edu).

# I. INTRODUCTION

Economists have long been aware that social networks may have a significant role to play in explaining economic outcomes. One challenge for empirical researchers has been to find measures of connectedness or social interaction that are sufficiently subtle to test nontrivial predictions of theoretical models. A second and perhaps more serious challenge has been to find sources of exogenous variation in social networks so that causal inferences may be drawn.[1] This paper attempts to address these challenges using network data from the National Longitudinal Survey of Adolescent Health (Add Health). The paper estimates the effect of connectedness (as measured by friendship ties in middle and secondary school) on education attainment and the relationship between racial heterogeneity and connectedness.

The peer effects literature provides a point of departure. While most recent studies find that an influx of high performing peers improves an individual's performance on standardized tests (Betts, Zau, and Rice [2003]; Hoxby [2000]; Hanushek, Rivkin, Markman, and Kain [2002]), the specific mechanism has not been clearly identified. How do peers influence one another? Do they teach each other? Do they imitate each other's effort choices? Do they create production externalities by disrupting the attention of others? Putting aside the question of the mechanism, one could also ask precisely which students influence which other students. Though relevant to policy implications, these questions have been difficult to answer in the empirical literature. The standard technique in recent studies has been to estimate the mean response by individuals to small exogenous changes (e.g., from students migrating in and out of specific schools, or from natural variations in gender composition by cohort) in the mean performance of their peers. There are indications, however, that mean peer performance does not tell the whole story. Angrist and Lang [2002] find that busing minority students into Boston public schools did not significantly alter the performance of white peers in the receiving schools but that the

[1] See Durlauf[2002], Manski[2000], and Moffitt[2001] for useful critiques of empirical practices in the social interactions literature.

performance of minority third graders did appear to be affected. Hanushek, Rivkin, and Kain [2004] find that a higher proportion of black students has an adverse effect on the performance of black students in Texas schools. The effect of racial composition appears larger for high achieving blacks than for previously low achieving black students or whites. All peers, it would seem, are not equal. There exists evidence, in short, that the degree of similarity between a student and a set of peers may determine the magnitude of his or her response, or even the direction of that response, to peer behavior. A number of current policies derive from the assumption that adding high-performing peers to low-performing schools or classrooms will benefit low-performers, unambiguously. This may not be the case.

The Add Health dataset contains detailed information about explicit friendship networks. A number of previous studies have used self-reported measures of "trust" as a proxy for connectedness or have relied on cross-city or cross-country variation in measures of trust or connectedness for identification.[2] Using the Add Health data, it is possible to see the individual student's friendship network when she was in school, and to thus to analyze how these may explain skill-related outcomes 7 years later. Secondly, it is possible to analyze correlations between school or grade-cohort characteristics and observed ties from individual to individual. This allows for a more micro-based approach to connectedness than has been possible in cross-region studies. New questions may be posed, strict quantitative methods applied, and the skill acquisition decision problem addressed from an uncommon perspective.

Results include 2 main findings: 1) Being part of a more connected cohort within a given secondary or middle school is associated with significantly higher years of schooling attained and higher probability of having attended college, 7 years later: 2) Being part of a more racially diverse grade cohort, within a given school, is associated with significantly lower levels of

[2] This is sometimes called the "social capital" literature. For purposes of clarity, I have avoided the term "social capital" in the present analysis. As is generally recognized, the term has different meanings in different settings. "Connectedness," by which I mean the size of the individual's friendship network, captures with greater specificity the relevant notion here.

connectedness. The paper is organized as follows: Section I motivates the inquiry; Section II presents a model of skill acquisition; Section III describes the data and explores racial heterogeneity as a determinant of connectedness; Section IV analyzes empirically the relationship between connectedness and long run skill acquisition; and Section V summarizes and draws conclusions.

## II. CONCEPTUAL FRAMEWORK

### A. A Model of Skill Acquisition Decisions

The analysis focuses on education as consumption good. In the standard model, forward-looking agents acquire skills as long as the opportunity cost of lost wages is offset by the increased future wages made possible by the incremental human capital acquired. Myopic agents, in contrast, acquire skills as long as the immediate utility they derive from doing so exceeds the utility derived from some outside option (e.g., dropping out). There is evidence that adolescents are less patient than adult agents.[3] If this is the case, then immediate consumption benefits may explain effort choices in skill acquisition as well or better than projected lifetime earnings. Consider a model in which agents value both academic rewards and the social rewards of school attendance, and individuals vary in their academic and social abilities. Agents maximize:

$$
\begin{aligned}
&\text{Max}_{e_a, e_s} U(A, S) \\
&\text{s.t. } e_a + e_s = 1 \\
&e_a \geq 0, e_s \geq 0 \qquad . \\
&A = f(e_a, \theta_a) \\
&S = g(e_s, \theta_s)
\end{aligned}
\tag{1}
$$

$A$ is the utility due to academic accomplishment. $S$ captures utility associated with social interactions. Effort may be directed toward production of academic accomplishment, $e_a$, or toward the production of social rewards, $e_s$. Agents are heterogeneous in their ability endowments, $\theta_a$, $\theta_s$, and production increases with effort input ( $f_{e_a} > 0, g_{e_s} > 0$ ). Here, $\theta_a$ and $\theta_s$

[3] See Bowdoin [2002], Nurma [1991].

raise the agent's production for the (internally produced) commodities *A* and *S*, respectively ($f_{\theta_a} > 0, g_{\theta_s} > 0$). If there is an interior solution, the utility gained from a marginal unit of effort devoted to producing the academic consumption good must equal the marginal benefit associated with effort directed toward producing the social consumption good:

(2) $$U_A f_{e_a} = U_S g_{e_s} .$$

$\underline{U}$ is the utility derived from the outside option—dropping out of school and taking a low skill job. (In general, $\underline{U}$ may depend on $\theta_a$, $\theta_s$, as abilities and tastes for activities in school may relate to abilities and tastes in the workplace.)

Figures IA and IB show the feasible consumption set and indifference curves for high-$\theta_s$ agents and high-$\theta_a$ agents respectively. The figure depicts a situation in which academic utility and social utility are perfect substitutes. (Linear utility in the figure is for simplicity of exposition.) One could imagine a change in the composition of schoolmates that altered the opportunity cost of making friends or of academic rewards. The usual assumption in peer effects literature is that high-endowment classmates make it easier for teachers to teach effectively and for low-endowment students to learn. An influx of new students with high parental education endowments (high $\theta_a$, as a result) would raise $f_{e_a}$ at every level of effort for their peers. (Marginal gains in academic accomplishment require less effort.) Evidence in Section III will indicate that students may acquire friends more easily when the school or grade-cohort population is more homogeneous. An influx of new students that reduces homogeneity might then lower $g_{e_s}$, as more effort may be required to make or to keep friends, or to navigate a factionalized environment. It is not hard to imagine both these changes occurring simultaneously, as depicted in Figure IC and ID. The marginal rate of transformation of academic utility for social utility falls. As drawn in IC, the decrease in $g_{e_s} / f_{e_a}$ lowers total utility for the low-$\theta_a$ individual to an amount below the level of the outside option, $\underline{U}$. Low-$\theta_a$ types (who would have stayed in school initially) drop out. As

drawn in Figure ID, high-$\theta_a$ types who initially would have dropped out experience an increase in utility and stay in school.[4]

It is possible to draw the production sets so that these outcomes do not obtain. Stronger assumptions on functional forms are needed to generate the effect unambiguously. If $c_1\theta_a$ and $c_2\theta_s$ are the constant marginal products $f'$ and $g'$, respectively, then the production set of internally produced commodities is non-convex and agents will specialize (in the case that utility is linear). Here $c_1$ and $c_2$ are meant to be suggestive of institutional factors that influence the ease with which academic and social utility are produced. In Figure IE and IF, $f = e_a c_1 \theta_a + d_1$ and $g = e_s c_2 \theta_s + d_2$. Because the starting point is an all-social utility corner solution in Figure IE, a marginal increase in $c_1$ (which increases the marginal product of effort for academic utility), will not cause substitution into academic effort and will not increase utility, whereas a marginal decrease in $c_2$ (which decreases the marginal product of effort for social utility) will reduce utility, unambiguously. High-$\theta_s$, low-$\theta_a$ types, in this example, are made worse off when $c_1$ rises and $c_2$ falls, while high-$\theta_a$, low-$\theta_s$ types (Figure IF) are made better off. The assumptions used in this example would seem rather strong; however, linear utility and constant marginal products were not strictly necessary. More generally, non-convexity of the production possibility set may lead to specialization by some agents, if there is strong substitutability between $A$ and $S$.[5]

[4] One could argue that having higher social ability raises rewards both at school and in the workplace (i.e., that $\underline{U}$ varies positively with $\theta_s$). An agent with high social ability would not necessarily then be more likely to stay in school. However, any exogenous change that made the individual more productive at acquiring friends in school while leaving unchanged the individual's capacity to acquire friends out of school would increase the likelihood that the individual stayed in school. It is institutional changes of this type that motivate the model and the empirical investigation.

[5] It is also possible that $A$ and $S$ could be complements. Academic accomplishment could make one popular. Social accomplishment could boost academic achievement by giving the agent friends from whom to receive tutoring. But because time must ultimately be divided between social and academic activities, and because of the large body of sociological evidence on identity and social groups in school settings, the account here will emphasize results that obtain if $A$ and $S$ are substitutes. Sociological evidence would seem to imply that agents face a choice between defining themselves as academic utility producers and social utility producers, as peers impose punishments for effort choices that deviate from group norms associated with these types. See Akerlof and Kranton [2002] for a summary of this evidence.

B. Stylized Predictions

One measure of *S* will be the number of friends the individual nominates or is nominated by—proxied in some settings, by the number of friendship links in a school or grade cohort. If production of *S* rises, holding *A* constant, utility stays the same or rises. (This holds for convex or non-convex production sets). If utility rises then the probability that the individual stays in school also rises.[6] Several stylized predictions, suggested by the modeling exercise, can be taken to the data:

1) Agents with higher *S* will stay in school longer (controlling for *A*). In addition, if "seat time" confers benefits or skills independent of academic performance, these agents should have labor market outcomes indicative of higher skill.[7]

2) If greater connectedness in an institution increases $g_{e_s}$ (i.e., reduces the effort required to produce social utility), then agents in connected institutions should stay in school longer.[8]

3) If there is specialization, production of *S* will generally determine the utility of low-$\theta_a$ types. Factors that hinder the production of *A* will not generally alter utility or influence the decision to stay in school. (This prediction is more fragile, as described above.)

## III. DATA AND CONNECTEDNESS

[6] This assumes that observed *S* is not positively correlated with the value of the outside option. If a high observed *S* indicates that an indivdual forms friendships easily, then this could indicate the individual would derive high utility from low-skilled work, as well (social interactions being a component of most work settings) and raise the value of the outside option. I assume here that the marginal beneft of increased social ability in school (where the agent would typically be exposed to hundreds of potential friends, long periods of time devoted solely to socializing, dances, proms, athletic events, and numerous clubs and organizations) is larger than in a low-skill job.

[7] Empirical work by Bowles, Gintis, and Osbourne [2001] suggests that schooling may indeed confer skills valued by the labor market that are not explained by measures of academic performance or cognitive skill.

[8] Some agents in some situations—those occupying an all academic effort corner solution—are unaffected. The prediction applies to agents at an interior optimum or an all social-effort corner solution.

A. Data

The Add Health survey, conducted by the Carolina Population Center from 1994 to 2002, consists of data on adolescents in 132 schools across the country, grades 7-12. The in-school portion of the Wave 1 survey, conducted in 1994-1995, contains cross-section data on about 90,000 adolescents. Also, school administrators filled out questionnaires describing characteristics and policies of the schools in the sample. A subset of the initial sample, about 20,000 subjects, was selected for the in-home portion of Wave I. This second and more extensive set of interviews with students and parents took place in 1994-1995. The Carolina Population Center interviewed the in-home subjects again in 1995-1996 (Wave II), and again in 2001-2002 (Wave III). When appropriate weights and cluster coefficients are used, regressions on data from each of the surveys, or from merged samples, yield results representative of the U.S. population.[9] Wave III measures of skill acquisition include years of schooling, college attendance and labor force participation about 7 years after the original Wave I surveys. These will be the dependent variables in the regressions that follow.

The most interesting aspect of the Add Health Survey, for the purpose at hand, is the data on friendship networks. Respondents in the in-school survey nominated up to 10 friends from the school roster, 5 male and 5 female.[10] The analysis here will be limited to the 113 schools in which at least 50% of the total student population filled out in-school questionnaires. For most of these schools, 75% or more of the entire student population filled out questionnaires. Thus, it is possible to see the general size and structure of friendship networks in these institutions. Figure II.A depicts as a directed graph the friendship network for one such school. UCInet and Netdraw were used to create this diagram.[11] Netdraw adjusts the shape of the diagram so that the physical distance between the nodes, as displayed, rises with the number of links on the shortest path

[9] In the following calculations, observations are weighted and errors are clustered at the school level—as recommended by designers of the Add Health survey. See Chantala and Tabor[1999].
[10] Some nominees were from a sister school. All school connectedness measures were constructed from same school nominations only.
[11] See Borgatti, Everett, and Freeman[2002].

between the nodes. Nodes far apart on the graph tend to have higher degrees of separation. Clustering of nodes in the diagram, then, indicates clustering in the network. In Figure II.A, the differing shapes of nodes represent different races. The startling feature of this graph is the extreme clustering of connections along racial lines. Very few links connect members of different racial enclaves. Though this is an extreme case, it offers a first suggestion that mean responses to mean peer performance could be misleading.

B. Connectedness

The argument presented in Section II was that connectedness lowers the effort cost of acquiring skill. Sociologists have constructed numerous measures of network centrality and connectedness. I focus here on several of the simplest: An agent's connectedness is the number of friends she has. Three different measures will be used. The in-degree of a student is the number of friendship nominations she received from other students in the school. In Figure II, a student's in-degree is visible as the number of arrows pointed toward her node. A student's out-degree is the number of students she nominated as friends. This is the number of arrows from her node to other nodes. The main measure of an individual's connectedness used in the analysis that follows is the number of friendship links with respect to which the individual is either the sender or receiver, the sum of in-degree and out-degree.[12]

In Section IV, "links" or connections will be the regressors of interest. A first strategy is to regress long–run education and employment outcomes directly on individual connectedness, as measured by links to and from other agents.

[12] If the agent derives utility primarily from friends with whom she interacts on a regular basis, then these would seem the appropriate measures. The emphasis here is on immediate utility, rather than on the flow of information across the structure of a network. Other measures of centrality and connectedness capture network characteristics relevant to the latter by counting friends of friends across multiple degrees of separation and/or weighting well-connected friends more heavily than less-connected friends. (See Wasserman and Faust[2000] for definitions and descriptions.) The direction of the sign on connectedness in the following regressions is robust across alternative measures of connectedness, though precision of the estimates varies. (Regressions featuring alternative measures of connectedness are available from the author upon request.)

1) $$Y_{ij} = \beta_1^1 C_i + \beta_2^1 X_{ij} + \beta_3^1 S_j + \beta_4^1 R_j + \varepsilon_{ij}.$$

Here, *i* indexes individuals, *j* indexes schools, $C_i$ is individual connections, $X_{ij}$ is a vector of individual characteristics, $S_j$ is a vector of school characteristics, $R_j$ is a vector of region, state, or city characteristics, and $\varepsilon_{ij}$ is the individual-specific error term. The potential for unobserved individual heterogeneity motivates a second specification:

2) $$Y_{ij} = \beta_1^2 C_{-i(j)} + \beta_2^2 X_{ij} + \beta_3^2 S_j + \beta_4^2 R_j + \varepsilon_{ij}.$$

The regressor of interest, $C_{-i(j)}$ (the construction of which will be described in Section IV), is the connectedness of the school, absent the individual's ties. A final (preferred) specification controls for unobserved school and neighborhood heterogeneity by including school dummy variables, $D_j$, characteristics of grade-level peers, $G_g$, and by defining connectedness at the grade level,

3) $$Y_{ijg} = \beta_1^3 C_{-i(g)} + \beta_2^3 X_{ijg} + \beta_5^3 G_g + \sum_j \alpha_j D_j + \varepsilon_{ijg},$$

The interpretation of $\beta_1^3$, the coefficient on grade-cohort connectedness, and tests for possible remaining sources of endogeneity will be discussed in detail Section IV.

Table I shows descriptive statistics of measures that will be used as right-hand variables in regressions of long-run skill on friendship ties. The average number of links to and from an individual in the sample is 8.74. The lower portion of Table I contains measures from the Add Health dataset that will be used as control variables. Covariates at the individual level include age, grade, sex, parental education, race, grade point average, and (for the smaller In-home sample) parental income, score on the Add Health vocabulary test, and average distance from the student's residence to the residences of schoolmates.[13] The following section motivates a focus

[13] Students from remote neighborhoods may be less apt to make friends in school. See the discussion in Section IV.

on measures of heterogeneity in discussions of connectedness, in particular, racial heterogeneity at the school level. The heterogeneity variable used here is a Herfindahl index subtracted from 1:

$$H_j = 1 - \sum_k r_{jk}^2 \tag{4}$$

where $r_{jk}$ is the fraction of the population of race $k$ in school $j$ and the racial categories are White, Black, Asian, American Indian, and Hispanic. Covariates at the school level also include school size, average class size, proportion male, PTA participation rates, percent new teachers, a breakdown of the school's population by race, average parental education, standard deviation of parental education, and average GPA. Covariates at the county level include racial heterogeneity (a Herfindahl index, as for schools), median income, standard deviation of income, and crime rate. State minimum wage may be a measure of the attractiveness of the outside option (dropping out of school) and has been included as well.

## IV. DETERMINANTS OF CONNECTEDNESS

### A. Connectedness and Heterogeneity

The motivation for exploring a possible relationship between connectedness and long-run outcomes would be immediately clear if connectedness is driven in part by policy. It is worth pausing, then, to identify institutional factors that may be determinants of connectedness. Moreover, there exists a large and intriguing literature that examines determinants of trust and connectedness in communities. One purpose of the present inquiry is to contribute to this growing literature—and possibly to fill a gap.

Much existing research work suggests that racial heterogeneity influences connectedness. Glaeser, Laibson, Scheinkman, and Soutter [2000] offer evidence from experiments using Harvard undergraduates that racial differences reduce trust and trustworthy behavior in games that yield a social surplus for cooperation. The same relationship has been documented in adult communities. Numerous researchers, including Poterba [1997], Alesina, Baquir, and Easterly

[1999], and Goldin and Katz [1999] find racial heterogeneity to be associated with lower provision of public goods. Empirical work by Alesina and La Ferrara [2002] indicates that levels of trust and cooperation in U.S. communities are highest in racially homogeneous neighborhoods. The evidence, however, derives mainly from cross-region and cross-country studies. Skeptics suggest that unobserved region or country-specific factors may have been confounded with ethnolinguistic fractionalization. What has been needed, arguably, is a systematic study of heterogeneity and connectedness using field data at the micro-level.

A first pass at the data, Column 1 of Table II shows results of an OLS regression of individual connectedness on individual, school, and neighborhood covariates, including school-level racial heterogeneity, $H_j$: School heterogeneity is negatively correlated with connectedness and significant at the 5 percent level. After school-level heterogeneity has been taken into account, however, increased county-level heterogeneity appears to be associated with higher connectedness. In addition, a high concentration of males appears bad for connectedness, whereas high parental education is associated with increased connectedness. Interestingly, school racial heterogeneity appears not to be a proxy for parental endowment heterogeneity or social class heterogeneity. Standard deviation of parental education (at the school level) appears to be marginally positively correlated with individual connectedness.

Figure III shows a scatterplot of number of friendship ties and school racial heterogeneity. The pattern shown in the regressions is clearly visible in the raw data. To contrast different types of schools close up, compare the directed graphs in Figures II.A and II.B. School 77 in Figure II.A—high in heterogeneity and low in connectedness—resides in the lower right portion of the scatterplot in Figure III, whereas school 146 in Figure II.B is from the upper left portion. School 77, with a heterogeneity index of .68, is more diverse than school 146, at .37. The number of friends per student in school 77 is 5.9, whereas in school 146, students have an average of 8.74 friends each, nearly 50% higher.

It may be that different subgroups respond in different ways to heterogeneity. In particular, one might suspect that heterogeneity leads to fewer ties only for white students, as they are often the majority group. This does not appear to be the case. Columns 2, 3, and 4 of Table II display regression results for several subsamples of the population. When the sample is restricted to blacks, the negative correlation between heterogeneity and connectedness persists; for whites, the coefficient remains negative, but is small in absolute value and imprecisely estimated.[14] One might suspect that heterogeneity helps those with low parental endowments and harms those with high parental endowments. This also appears not to be the case. The negative correlation between racial heterogeneity and connectedness persists for the sample of respondents with low parental education.

Neighborhood effects and unobserved school heterogeneity may be driving the results above. Advantaged families may select into homogeneous schools. The main set of regressions for this section is summarized in Columns 5, 6, and 7 of Table II. To account for neighborhood and school-specific effects, I include school dummies in these regressions and analyze individual connectedness as a function of racial heterogeneity at the grade level. Columns 6 and 7 decompose the total links variable into out-degree and in-degree, respectively.[15] Grade cohort racial heterogeneity is negatively correlated with connectedness at significant levels for all measures. Being part of a more racially heterogeneous grade cohort within a given school is associated with having significantly fewer friends.

[14] American Indians are a small subpopulation of the full sample. Thus, the large coefficient on percent American Indian for the white subsample in Column 3 might seem peculiar. When American Indians are dropped from the school racial heterogeneity index, the coefficient on school heterogeneity remains negative and insignificant. Results in other regressions are robust in magnitude and statistical significance to the use of this alternative heterogeneity index.

[15] Because some nominations may go from members of one grade to members of another grade, it could be that regressions using out-degree would tell a different story than regressions using in-degree.

B. Discussion.

Traditional peer effects—to the extent that they can be captured by grade-level GPA, grade-level parental education, grade-level dispersion of parental education, and grade-level gender composition—do not appear to drive the relationship between racial heterogeneity and connectedness in the school fixed effects regressions. Though boys appear to have fewer ties than girls, the large negative coefficient on percent male in the baseline regression becomes marginally positive in the school fixed effects specifications, suggesting that there may be idiosyncratic unobserved factors in schools with a high percentage of males.

In the school fixed effects regressions, I assume that changes in racial heterogeneity at the grade cohort level are driven by random variations in the demographic composition of these cohorts. It remains possible that agents select into and out of specific grades within a school, based on the racial heterogeneity of the grades—in which case the estimates in these regressions might capture the effects of selection. Arguably, however, selecting into and out of specific grades in this manner would seem difficult: Fluctuations in racial heterogeneity between grade cohorts within a given school would seem hard to predict.[16]

Recent research in sociology has used the Add Health dataset to analyze friendship networks and race, and the research complements the findings here (Joyner and Cao, 2000, Moody, 2001, Quinlan and Campbell, 2003.) The present investigation differs from existing work in two important ways. Firstly, the primary purpose of these existing studies has been to describe and understand the determinants of cross-race friendships relative to same-race friendships—and thus to discern whether racial diversity at the school-level implies racial diversity in friendship networks. Because of this focus, they do not analyze, in a direct way, the influence of racial

[16] One could argue that parents of children with high social endowments choose to withdraw their children from a grade if the grade-cohort proves too heterogeneous. Even so, the costs of switching after learning grade-specific heterogeneity may be nontrivial.

heterogeneity on total number of friends or connectedness.[17, 18] The present investigation is concerned with connections as a source of utility (or of lower disutility of acquiring human capital). I do not assume differences in the impact of between-race and within-race friendship ties on immediate utility. Secondly, the studies listed above lack identification strategies to account for school-level unobservables—arguably an important set of factors, the effects of which could easily be confounded with the effect of racial heterogeneity.

In the economics literature, a number of studies have found a relationship between racial heterogeneity and low levels of trust and connectedness. In many existing studies, it is difficult to discern whether unobserved characteristics at the country or regional level have been confounded with racial heterogeneity. Studies based on cross-region and cross-country variation in racial heterogeneity would seem more susceptible to criticisms of this type than would a study based on within-school variation in racial heterogeneity between grade cohorts. The analysis here attacks the problem at a much lower level of aggregation. As such, it may have advantages. Moreover, the findings here—if indicative of a causal channel—may have policy relevance, particularly in light of evidence on connectedness and human capital acquisition in the next section.

## IV. CONNECTEDNESS AND LONG-RUN OUTCOMES

### A. School-Level Results

Is connectedness correlated with long-run skill acquisition? Table III, a first pass, displays findings from regressions of long-run skill-related outcomes on individual connectedness

[17] Moody [2001], for example, finds that the odds ratio for a same-race friendship nomination at first increases with the racial heterogeneity of the school, then decreases modestly at the highest levels of heterogeneity. Moody's argument is that if racially diverse schools feature students who interact most often with same-race friends, segregation remains: "The lived experience of students in such settings is of racial division, not integration." The purpose is to discern whether diverse schools lead to diverse friendships. Quinlan and Campbell (2003) extend the analysis to Hispanic and Asian students and examine cross-race ties over different pairings of races.

[18] There is also research on race and friendship relations in the economics literature. Fryer and Torrelli (2005) focus on same-race friendships. They offer evidence from Add Health friendship data of large racial differences in the relationship between popularity and GPA. Blacks and Hispanics appear not to be rewarded as much with popularity for higher GPAs as whites are, and appear to lose popularity when their GPAs are in the higher regions of the distribution.

("links") and covariates. I use total number of friendship ties as the measure of connectedness in this and subsequent regressions. (When In-degree or Out-degree is used, results are very similar). Results from OLS and Probit regressions of years of schooling, college participation and labor force participation[19] on connectedness and controls appear in columns 1, 2 and 3 of Table III. Estimates of the coefficient on links are positive and significant. Other coefficients have the expected signs; higher family income leads to higher skill-related outcomes, as do higher parental education and higher scores on the Add Health vocabulary test.

If a person who reports having numerous friends defines friendship in a less restrictive way than does a person who reports few friends, the willingness to nominate friends could be associated with advantageous personal traits such as optimism and confidence. These traits may drive the observed correlation with measures of long-run skill. To test for this, I ran the regressions using In-degree—the number of friendship nominations received—as the dependent variable. The results were qualitatively very similar—a strong indication that subjective definitions of friendship do not drive the correlation.

It could also be that innately gregarious or charismatic individuals experience long-run outcomes indicative of higher skills because social ability is itself a skill.[20] In Columns 4,5, and 6, school-level connectedness replaces individual connectedness as the regressor of interest. Individuals attending a friendlier institution might derive more immediate utility and stay in school longer, regardless of their individual endowments of gregariousness or charisma.

Consider a school with $n$ students. One measure of school connectedness for individual $i$ in school $j$ is

$$C_{-i(j)} = \frac{\sum_{k \neq i, l \neq i} X_{kl}}{10(n-1)} \tag{4}$$

[19] Some respondents were still in school at the time of the Wave III surveys. Labor force participation regressions include only those respondents who were not still in school at the time of the Wave III surveys.
[20] If (in the workplace) social skill were a complement to education, then individuals with higher initial endowments of social skill would choose more schooling because the return to schooling would be higher for these individuals.

where $X$ is an $n$ X $n$ matrix, $X_{kl}$=1 if $k$ nominated $l$ as a friend and 0 otherwise. The numerator is the number of links in the network minus the number of links to or from node $i$. Because students may nominate a maximum of 10 friends, the maximum number of links in the network is 10$n$. If we imagine removing one student from the network, the maximum number of links would be 10($n$-1).[21] $C_i$, then, measures the connectedness of the network, absent $i$'s links.[22] Multiplying this measure by 10 yields a measure, $\widetilde{C}_{-i(j)}$, with a more straightforward interpretation: the average number of links per student (absent $i$'s links) for $i$'s schoolmates. Columns 4, 5, and 6 in Table III use this second measure.

Clearly, school-level unobservables could be driving the results in Table III. Table IV addresses this concern by including school-level dummy variables for all specifications. Further, in Columns 4, 5, and 6 of Table IV, grade-cohort connectedness is the regressor of interest (and grade-level control variables replace school-level controls.) Though the effect of connectedness on labor participation in Column 3 goes away in Column 6, the positive and statistically significant correlation between connectedness and education outcomes persists across specifications: Having been part of more connected grade cohort within a given school is associated with higher levels of schooling attained and greater probabilities of attending college in the long run. This would seem to rule out school and neighborhood specific unobservables (though not grade-specific unobservables) as the source of correlation between connectedness and observed education outcomes.

The magnitudes of the connectedness coefficients would seem economically significant as well as statistically significant. Table IV, Columns 4, 5—the preferred specifications—indicate

[21] The maximum of 10 allowed friendship nominations was not so low as to make the data uninformative. Only about 3% of the respondents nominated 10 friends. For the vast majority of respondents, the restriction did not bind.

[22] One could argue that $C$ may be biased toward large schools. Agents attending large schools, having more friends from which to choose, might be expected to have more friends, on average. I have included school size in the regressions that follow to account for this possibility. Moreover, the data indicate that students in large schools appear not to have significantly more friends than students from small schools. (See Table III).

that an increase of 1 link in the average links per student in *i*'s grade cohort is associated (about 7 years later) with an increase of .13 years of schooling acquired by *i* and an increase of 2 percentage points in the probability that *i* attended college. (Here, the latter is the weighted population mean of the estimated marginal effects in the "college" probit regression.) One standard deviation of grade cohort connectedness is about 2.1. Thus, the association in Columns 5, if causal, implies that a 1-standard-deviation increase in grade-cohort connectedness raises average years of schooling by a quarter year and the probability of college attendance by 4 percentage points. These effects would seem large enough to be of interest to parents, policy-makers, and median voters.

B. Robustness and Alternative Explanations

*1. Instrumental Variables*

An alternative strategy would have been to use grade-cohort connectedness as an instrument for individual connectedness. Indeed, the coefficient on individual connectedness could then be interpreted as the marginal increase in education outcomes associated with having 1 more friendship tie. Estimated coefficients on individual connectedness in IV regressions are close to the analogous coefficients on grade level connectedness in Columns 4, 5, and 6, in magnitude and statistical significance.[23] There are reasons, though, to doubt the instrument's validity. The individual may derive greater utility from being part of a more connected cohort, even if the individual does not acquire more friends herself.[24] If this were the case, the instrument would be correlated in a direct way with the outcome variables. The assumption that having more friends is the only utility-enhancing aspect of being in a more connected

[23] Results of IV regressions are available from the author upon request. The strong correlation between grade level connectedness and individual connectedness in the first stage is built into these measures. Thus, the effect of the first stage is to "scale" the coefficient on the instrument by a factor close to 1.
[24] For example, an awareness of tensions between one's schoolmates could create unease and reduce an individual's social utility.

environment could be problematic [25]—and would seem unnecessary for interpreting results. The argument is that being part of a more connected cohort raises social utility, S. I have emphasized regressions in Table IV that leave open the channel through which the influence works.

*2. Simultaneity*

Agent *i*'s friend-making behavior may influence others in her school or grade cohort, causing her friends to form more friendships with third parties. If this were the case then *i*'s own friendliness would explain, in part, the connectedness of *i*'s school or cohort (excluding *i*'s links). This is akin to the "reflection problem," as formalized in Manski[1993]. To mitigate this problem, I excluded small schools from the sample. If there are only a few students in *i*'s school then *i*'s charisma could plausibly influence the average number of friendships *i*'s schoolmates form with third parties. If there are numerous students in *i*'s school, then it would seem less plausible that *i*'s friendliness influences significantly the number of ties between *i*'s schoolmates and third parties. Results were robust to the exclusion of small schools and schools with small grade cohorts. Moreover, the coefficient on grade cohort connectedness did not fall as small schools were excluded.

*3. Neighborhood Friends*

School friends and neighborhood friends need not be identical, as a given school draws students from a multiplicity of neighborhoods. The Add Health survey data contain only school friends. If neighborhood characteristics influencing long-run outcomes are correlated with an individual's propensity to form neighborhood ties (as opposed to school friendship ties), then the coefficients on connectedness in Table IV could be biased upwards. A story about disadvantaged

[25] Racial heterogeneity of the grade cohort might seem a plausible instrument for connectedness, given the association between connectedness and heterogeneity in Table II. However, the Wave III sample is much smaller than the Wave I in-school sample. The coefficient on racial heterogeneity is imprecisely estimated in the smaller sample. Thus, well-known problems associated with weak instruments argue against using racial heterogeneity as an instrument for connectedness.

students being bused to distant schools captures the intuition. If disadvantaged students make fewer friends at school because they live far away from their schoolmates, then the number of friendship ties could be a proxy for unobserved disadvantage. The Add Health data contain information on spatial relationships between respondents. Figure IV shows the relative locations of residences for students attending School 77. From this data, I computed the average distance between each student and her schoolmates. If distance captures in part the student's propensity to form in-school friendships as opposed to neighborhood ties, and if there exists a systematic correlation between distance and disadvantage, then including distance should lower the coefficient on individual connections. Regressions in Tables III and IV already include the distance measure and the coefficient is negative in all specifications.

*4. Peer Effects and Cognitive Ability*

There are compelling reasons to believe that grade-cohort connectedness influences education outcomes through a channel involving traditional peer effects—and it is worth attempting to distinguish this channel from the social utility channel proposed in the model. Friendship data in the Add Health survey are based on students filling out a questionnaire in which they write in names of their friends and find codes for those names on a list of ID numbers for students in their school. One might worry that students who are better at taking tests would also be better at following instructions, more likely to fill out the friendship questionnaire more completely—and thus would nominate more friends. In such a case, friendships would be an indirect measure of cognitive skill, and correlated with education attainment. Columns 4, 5 of Table IV—the main results in this section—regress individual education outcomes on grade level connectedness. Does the positive coefficient on connectedness, then, reflect a traditional peer effect—namely, that students with grade-level peers of higher cognitive ability will have outcomes indicative of higher cognitive skill?

The regressions already include peer characteristics as covariates, but one could go further. Students in the Wave I sample completed the Add health vocabulary test. If the connectedness of an individual's grade cohort is primarily a proxy for cognitive skill, then one would expect a positive correlation between test scores on this vocabulary exam and the connectedness of the grade cohort. Similarly one would expect a positive correlation between the individual's GPA and the connectedness of the grade cohort. School fixed effects regressions of Add Health vocabulary test score on grade cohort connectedness (with or without the full set of controls) yield negative and statistically insignificant estimates of the grade-cohort connectedness coefficient, as do similar regressions of individual GPA on grade cohort connectedness. Both are summarized below[26]:

AH Vocab = -.047 Links (grd) + school dummies
(.233)

GPA = -.023 Links (grd) + school dummies
(.020)

Grade cohort connectedness is not associated with higher individual test scores or with higher GPAs—a finding which would seem to cast doubt on the notion that connectedness of grade cohort is merely a proxy for cognitive skill.

*5. Peer Quality vs. Peer Quantity*

To further distinguish traditional peer effects from the effects of connectedness, I explore a setting in which the two effects would work in opposite directions. The concern about a possible confounding influence of traditional peer effects in Table IV could be characterized as a distinction between peer quality and peer quantity. If the results in Table IV are driven by traditional peer effects, one would predict that having more low-performing peers would be worse than having fewer low-performing peers. In other words, if it is the quality of peers that drives

[26] These include only school dummies as controls. Results for regressions that include the full set of covariates are similar.

skill acquisition outcomes, rather than the quantity, then having more "bad" peers (a larger number of bad examples to imitate) would lower academic performance and skill-related outcomes. If, on the other hand, connectedness to one's schoolmates drives skill acquisition outcomes, then having more peers (even "bad" ones) raises social utility relative to the outside option and keeps one in school.[27]

Panel A of Table V restricts the Wave III sample to respondents whose nominated friends had an average GPA lower than the school's average GPA. These students, then, had low-performing peers. Column 1 indicates that having more friends of this type was associated with having earned a (marginally) lower GPA. This would be the expected peer effect (or the result of selection by low-effort types into low-performing peer groups, as described below.) Panel B, Column 1 indicates that if one's peers were high academic performers, having more peers increased one's own academic performance (or was a stronger indicator that one was a high performer to begin with).

Panel A, Columns 2 and 3, flesh out the story in a provocative way. Even if one's peers are **low performers**, having more of them is strongly associated with a **higher probability of finishing high school and going on to college**. (The same story holds when the dependent variable is years of education.) There would seem to be more than one set of motivations for staying in school. In the model of section II, students derive utility from internally produced social rewards. These raise the utility of education as consumption good, relative to the outside option, and increase the likelihood that the agent will continue on in school. The peer effect exerts a downward pressure on GPA, but increased utility associated with social rewards leads to higher levels of skill acquisition. The evidence would suggest that the connectedness effect dominates the peer effect.

[27] For the above interpretation to hold, it would need to be the case that social ability had a relatively minor effect on utility in the workplace, but a large effect on utility in school.

Moreover, it is likely that the peer effect is smaller in magnitude than might be inferred from Table V. Low-performing students select low-performing friends. Having more low-performing friends is then a stronger indicator that one was a low performer to begin with. Despite the effects of selection, students with more low-performing friends stay in school longer. Evidence suggests that the effects of connectedness dominate **both** the peer effect and the selection effect.

This section also tests the third prediction of Section II: Factors that hinder the production of academic rewards but facilitate the production of social rewards will not generally influence the decision to stay in school for types with low academic ability. The intuition is that agents with low academic ability would have been consuming more social rewards and fewer academic rewards in the first place. Thus, an increase in the effort cost of academic rewards does little harm, whereas a decrease in the effort cost of social rewards benefits them. In spite of (marginally) lower GPAs, agents with more low-performing friends consume more education, because social rewards are the dominant component of utility for low-performers. The pattern of empirical results in Table V, though somewhat counterintuitive, is consistent with the model.

*6. Grade-Specific Teacher Quality*

In addition to sharing grade-level peers, individuals in a grade cohort share a common set of teachers.[28] Perhaps talented teachers in a grade-year cohort cause students both to make more friends (or to fill in more names on the questionnaire) and to have higher educational attainment. In a world in which students and teachers did not move between schools, each student would eventually receive the "teacher quality" treatment for each grade in the school. Thus, this channel would seem to be driven primarily by movements of students and/or teachers in and out of schools or by changes in teacher quality over time. One may infer that between Wave I and Wave

[28] See Moffitt [2000] for an extended discussion of common environment and other issues relevant to empirical work on peer effects.

II, 12% of students in the sample (or less) left their school without having completed the grades offered in that school.[29] Unfortunately, there is no data available on teacher movements and grade-specific teacher quality.

Grade-specific unobservables cannot be ruled out as a possible confounding factor. Arguably, though, the teacher quality explanation, if true, would also suggest that social utility or connectedness (as facilitated by teachers) drives outcomes: "Teacher quality," here, would seem to cause students to claim more friends and to stay in school longer, but to do so without increasing their test scores or their GPAs.

## C. Discussion

Comparisons with existing research provide a sense of the magnitude of the coefficient on connectedness. A causal interpretation of the preferred Table IV regressions implies that an addition of 2 links to the average connectedness of grade-level peers would wipe out the education attainment gap between blacks and whites.[30] Is this an implausibly large effect? Related work suggests not. Ehrenberg and Rothstein(1994) find that students at historically black colleges were 9% to 29% more likely to graduate than similarly qualified black students at other 4-year colleges. And Constantine(1995) finds that they went on to earn higher wages.[31] (Moreover, the findings obtain despite lower average peer SAT scores at historically black colleges.) Though connectedness, as such, is not measured in these studies, one could infer that connectedness, cultural cohesiveness, or social utility at historically black colleges drives this seemingly large observed effect.

[29] This was based on attrition from the sample between Wave I and Wave II other than through finishing the final grade offered in a school, and equates to about 20% of the students who were in non-final grades. It may overstate student movement because other factors could have led to attrition. (Few students changed from one surveyed school to another surveyed school between the two surveys.)
[30] See Maxwell[1994].
[31] The estimated wage gap ranged from 11% to 38%.

Other evidence of large effects of connectedness on long-run outcomes comes from Galleoti and Mueller (2005), who use the Washington Longitudinal Study to analyze friendship relations in high schools as predictors of adult wages. After controlling for school-specific effects, cognitive ability, grade rank in class, and personality traits, they find that differences in friendship ties in high school predict large wage differences, 35 years later. Galleoti and Mueller's study is related to the present analysis, but differs in important ways. There is only one grade and cohort in the sample—and the students were in high school in 1957. A maximum of 3 friends could be nominated. Information on the full school or grade network was unavailable. Main findings pertained to "isolates" (students with no friends.) Though an advantage of the WLS is that it contains long-run outcomes 35 years after high school, the dataset would seem to limit potential strategies to distinguish the effect of individual endowments from the effect of institutional factors. Moreover, much has changed in 50 years—including the dramatic rise in the wage premium for educated workers.

### D. Magnet Schools and Education Policy

Magnet schools were intended, in part, to attract advantaged students to schools in disadvantaged areas so that positive peer effects would benefit disadvantaged students. The consequences of these policies with respect to connectedness may not have been fully appreciated. Anecdotal accounts indicate that advantaged students in magnet schools often do not mingle with disadvantaged students. Is there systematic evidence of lower connectedness at magnet schools?

Table VI shows that the average number of connections in the 15 magnet schools surveyed in the Add Health sample is 6.8. In some schools, students are assigned from several geographic areas in order to achieve a desired racial composition. Six schools reported using an assignment rule of this type (and 4 of these were also Magnet schools). Table VI shows the average number of connections in schools that assigned pupils to achieve racial balance was 6.2.

Both these numbers are a full standard deviation lower than the average of 8.9 connections per student found in the 108 other schools.

The evidence in Table VI is suggestive and thought-provoking, if not conclusive: Policies that move disparate groups of students between schools may lower connectedness. And if disconnected student populations raise the disutility of acquiring schooling, some policies of this type may have unintended consequences. Ultimately, there may be no good substitute for improving the schools students attend. More generally, it could be the case that women's colleges, historically black colleges, and other institutions that appeal to individuals with common cultural bonds may generate positive outcomes by harnessing connectedness.

## VI. SUMMARY AND CONCLUSION

The evidence indicates that an individual's educational attainment increases with the connectedness of her school. Moreover, being part of a more connected cohort within a given secondary or middle school is associated with significantly higher levels of schooling attained and higher probabilities of having attended college, 7 years later. Lastly, being part of a more racially heterogeneous cohort within a given school is associated with having significantly fewer friendships.

A research agenda that focuses on test scores and peer effects may lead policy makers to overlook the effects of connectedness on skill investment decisions, effects that may be quite large. Evidence in Section II showed that connectedness varies with school or grade cohort heterogeneity. The evidence is consistent with a growing body of research on heterogeneity in adult communities. This is not to suggest that people cannot learn to cooperate within diverse populations over the long haul, but it does suggest there may be hidden costs in the short to medium run. These costs may be worth knowing, so that they may be balanced against the clear,

long-run benefits of diversity—or so that institutions and policies that promote connectedness may be developed and tested.

Collection and analysis of network data in a controlled setting (with random school or classroom assignment) may be warranted. More generally, it is possible that the collection of network-type data in college or labor market settings, and the subsequent application of the economist's statistical toolkit to this data, could yield new and useful insights. These are subjects for future research.

References

Akerlof, George and Rachel Kranton, "Identity and Schooling: Some Lessons for the Economics of Education," *Journal of Economic Literature*, XL(2002), 1167-1201.

Alesina, Alberto, Reza Baqir, and William Easterly, "Public Goods and Ethnic Divisions," *Quarterly Journal of Economics*, CXIX(1999), 1243-1284.

Alesina, Alberto, and Eliana La Ferrara, "Ethnic Diversity and Economic Performance," NBER Working Paper 10313, 2004.

____ and ____, "Who Trust Others?" *Journal of Public Economics*, LXXXV(2002), 207-234.

Angrist, Joshua and Kevin Lang, "How Important are Peer Effects? Evidence from Boston's Metco Program," NBER Working Paper #9263, 2002.

Austin-Smith, David and Roland Fryer, "An Economic Analysis of 'Acting White'," *Quarterly Journal of Economics*, CXX(2005), 551-584.

Betts, Julian, Andrew Zau, and Lorien Rice, *Determinants of Student Achievement: New Evidence from San Diego* (San Francisco: Public Policy Institute of California, 2003).

Borgatti, Steve, Martin Everett, and Linton Freeman, *Ucinet 6.87 for Windows: Software for Social Network Analysis* (Harvard, MA: Analytic Technologies, 2002).

Bowdoin, Scott David, "A Comparison of Decision-Making By Adolescents and Adults," manuscript, Moore School of Business, University of South Carolina, 2002.

Bowles, Samuel, Herbert Gintis, and Melissa Osbourne, "Incentive-Enhancing Preferences: Personality, Behavior, and Earnings," *American Economic Review*, XCI(2001), pp.155-158.

Chantala, Kim and Joyce Tabor, "Strategies to Perform a Design-Based Analysis Using the Add Health Data," mimeo, Carolina Population Center, University of North Carolina at Chapel Hill, 1999.

Constantine, Jill, "The Effect of Attending Historically Black Colleges and Universities on Future Wages of Black Students," *Industrial and Labor Relations Review*, XLVIII(1995), 531-546.

Durlauf, Steven, "On the Empirics of Social Capital," *The Economic Journal*, Vol. CXII(2002), pp. 459-479.

Ehrenberg , Ronald and Donna Rothstein, "Do Historically Black Institutions of Higher Education Confer Unique Advantages on Black Students? An Initial Analysis," in *Choices and Consequences: Contemporary Policy Issues in Education,* R. Ehrenberg, ed. (Ithaca, N.Y.: ILR Press, 1994).

Fryer, Roland and Paul Torelli, "An Empirical Analysis of 'Acting White'," NBER Working Paper 11334, 2005

Galeotti, Andrea and Gerrit Mueller, "Friendship Relations in the School Class and Adult Economic Attainment," IZA Discussion Paper No. 1682, 2005.

Glaeser, Edward, David Laibson, Jose Scheinkman, and Christine Soutter, "Measuring Trust," *Quarterly Journal of Economics*, LXV(2000), 811-846.

Goldin, Claudia, and Lawrence Katz, "Human Capital and Social Capital: The Rise of Secondary Schools in America," *Journal of Interdisciplinary History*, XXIX(1999), 683-723.

Joyner, Kara, and Grace Kao, "School Racial Composition and Adolescent Racial Homophily," *Social Science Quarterly*, LXXXI(2000), 810–825.

Hanushek, Eric, John Kain, Jacob Markman, and Steven Rivkin, "Does Peer Ability Affect Student Achievement?" NBER Working Paper 8502, 2001.

Hanushek, Eric, John Kain and Steven Rivkin, "New Evidence about Brown v. Board of Education: The Complex Effects of School Racial Composition on Achievement," mimeo, Stanford, 2004.

Hoxby, Caroline, "The Effects of Class Size and Composition on Student Achievement: New Evidence from Natural Population Variation," *Quarterly Journal of Economics*, CXV(2000), 1239-1285.

Manski, Charles. 2000. "Economic Analysis of Social Interactions," *Journal of Economic Perspectives,* XIV*,* 115–136.

Manski, Charles, "Identification of Endogenous Social Effects: The Reflection Problem," *The Review of Economic Studies*, LX(1993), 531-542.

Maxwell, Nan, "The Effect of Black-White Wage Differences of Differences in the Quantity and Quality of Education," *Industrial and Labor Relations Review*, XLVII(1994), 249-264.

Moffitt, Robert, "Policy Interventions, Low-level Equilibria, and Social Interactions," in *Social Dynamics*, S. Durlauf and H.P. Young, eds. (Cambridge: MIT Press, 2001).

Moody, James, "Race, School Integration, and Friendship Segregation in America," *American Journal of Sociology*, CVII (2001), 679-716.

Nurmi, Jari-Erik, "How Do Adolescents See Their Future? A Review of the Development of Future Orientation and Planning," *Developmental Review*, XI(1991), 1-59.

Poterba, James, "Demographic Structure and the Political Economy of Public Education," *Journal of Policy Analysis and Management*, XVI(1997), 48-66.

Quinlan, Lincoln and Mary Campbell, "Beyond Black and White: The Present and Future of Multiracial Friendship Segregation,"*American Sociological Review*, LXVIII(2003), 540-566.

**Table I**
**Descriptive Statistics**

| | Mean | St. Dev. |
|---|---|---|
| Links | 8.74 | 5.68 |
| Out Degree | 4.37 | 3.05 |
| In Degree | 4.37 | 3.70 |
| | | |
| Age | 14.86 | 1.78 |
| Parental Ed | 13.01 | 2.37 |
| Parental Incom (1000s)[a] | 35.84 | 44.62 |
| AH Vocab Test[a] | 97.62 | 25.52 |
| Distance (ave. km, peers)[a] | 7.12 | 6.54 |
| GPA | 2.44 | 1.20 |
| White | 0.57 | 0.50 |
| Black | 0.20 | 0.40 |
| Asian | 0.04 | 0.19 |
| Amer. Ind. | 0.04 | 0.19 |
| Hisp | 0.12 | 0.33 |
| Male | 0.51 | 0.50 |
| | | |
| Grade 7 | 0.17 | 0.38 |
| Grade 8 | 0.17 | 0.37 |
| Grade 9 | 0.17 | 0.38 |
| Grade 10 | 0.17 | 0.37 |
| Grade 11 | 0.16 | 3.97 |
| Grade 12 | 0.16 | 0.37 |
| | | |
| Het. Race (sch) | 0.39 | 0.19 |
| Par. Ed. (sch) | 12.99 | 0.85 |
| Std. Dev. Par. Ed. (sch) | 2.08 | 0.27 |
| Class Size(sch) | 25.48 | 5.04 |
| Size (sch) | 994.60 | 747.89 |
| Prop. New Teachers(sch) | 0.99 | 0.15 |
| Prop. PTA Participation(sch) | 0.23 | 0.22 |
| GPA(sch) | 2.87 | 0.25 |
| Prop. Male(sch) | 0.50 | 0.06 |
| | | |
| Juv. Crime (per 100K, cnty) | 341.52 | 171.96 |
| Het. Race (cnty) | 0.29 | 0.18 |
| Med. Income (cnty: in 1000s) | 29216.65 | 7853.33 |
| Std. Dev. Income (cnty) | 28951.63 | 5343.01 |
| Min Wage (state) | 4.28 | 0.16 |
| Crime Rate (per k, state) | 5283.49 | 1157.04 |

[a]These variables available for In-home sample. All other covariates available for full sample.

**Table II**
**Determinants of Connectedness[a]**

| | **Baseline** | **Black** | **White** | **P. Ed<12** | **School Fixed Effects** | | |
|---|---|---|---|---|---|---|---|
| Dependent Variable | Links | Links | Links | Links | Links | Out | In |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Het. Race (sch or grd)[c] | -1.93** | -4.62*** | -.175 | -2.09* | -3.69** | -1.92** | -1.77* |
| | (.882) | (1.43) | (1.52) | (1.15) | (1.68) | (.822) | (.911) |
| Prop. Black (sch or grd)[c] | -2.03*** | -1.69* | -4.55*** | -2.22*** | 2.17 | 1.41 | .758 |
| | (.641) | (.947) | (.969) | (.841) | (2.2) | (1.12) | (1.14) |
| Prop. Asian(sch or grd)[c] | -4.4** | 3.2 | -9.73*** | -1.99 | 5.22** | 2.79** | 2.43* |
| | (2) | (3.68) | (2.61) | (2.26) | (2.47) | (1.28) | (1.31) |
| Prop. Amer. Indian (sch or grd)[c] | -3.11 | .655 | -22.7*** | .114 | 4.27 | 2.21 | 2.06 |
| | (6.27) | (8.34) | (6.51) | (4.71) | (3.79) | (2.02) | (2.05) |
| Prop. Hisp (sch or grd)[c] | -1.46 | 1.26 | -3.62 | -.178 | -.825 | -.0747 | -.751 |
| | (1.45) | (2.39) | (2.36) | (1.94) | (2.28) | (1.21) | (1.16) |
| Prop. Male (sch or grd)[c] | -8.58*** | -9.45*** | -9.35*** | -6.42** | -1.19** | -.717** | -.472 |
| | (1.53) | (3.02) | (1.82) | (2.59) | (.598) | (.303) | (.321) |
| Size (100s - sch or grd)[c] | -.0294 | -.000641 | -.00299 | -.0404 | .0604 | .0501 | .0103 |
| | (.0223) | (.0396) | (.0216) | (.0296) | (.152) | (.0714) | (.0848) |
| Ave GPA (sch or grd)[c] | .277 | .345 | .222 | .0952 | -.731* | -.451** | -.279 |
| | (.524) | (.905) | (.476) | (.721) | (.438) | (.204) | (.259) |
| Ave. Parent Ed. (sch or grd)[c] | -.129 | .155 | -.406 | -.202 | .49 | .288* | .202 |
| | (.232) | (.381) | (.257) | (.325) | (.303) | (.157) | (.156) |
| Std. Dev. Parent Ed. (sch or grd)[c] | .968 | -1.4 | .832 | 1.06 | -.126 | -.0297 | -.0963 |
| | (.689) | (1.58) | (.641) | (.891) | (.446) | (.218) | (.245) |
| Parental. Education (ind) | .214*** | .127*** | .273*** | .342*** | .212*** | .0745*** | .138*** |
| | (.0195) | (.0468) | (.0226) | (.0642) | (.0199) | (.00851) | (.0131) |
| GPA(ind) | .726*** | .415*** | .904*** | .649*** | .796*** | .349*** | .446*** |
| | (.0511) | (.0795) | (.064) | (.0946) | (.0505) | (.0268) | (.031) |
| Black (ind) | -.891*** | - | - | -.343 | -.921*** | -.585*** | -.336*** |
| | (.149) | | | (.24) | (.148) | (.0702) | (.0835) |
| Asian(ind) | -.783*** | - | - | -.886** | -.846*** | -.458*** | -.389*** |
| | (.19) | | | (.392) | (.187) | (.0956) | (.106) |
| Amer Indian (ind) | -.67*** | - | - | .194 | -.71*** | -.271*** | -.44*** |
| | (.154) | | | (.274) | (.149) | (.0854) | (.0826) |
| Hispanic(ind) | -.324** | - | - | .00867 | -.342*** | -.278*** | -.0644 |
| | (.126) | | | (.232) | (.125) | (.0714) | (.0676) |
| Male(ind) | -.604*** | -.339** | -.721*** | -.698*** | -.641*** | -.387*** | -.254*** |
| | (.0674) | (.164) | (.0891) | (.122) | (.0641) | (.0461) | (.042) |
| Het. Race (cnty) | 2.71* | 2.29 | 3.43** | 1.81 | - | - | - |
| | (1.45) | (2.32) | (1.42) | (1.5) | | | |
| Med. Income (cnty: in 1000s) | .0297 | .11* | .0171 | .0178 | - | - | - |
| | (.0391) | (.0583) | (.0393) | (.0457) | | | |
| Std. Dev. Income (cnty) | -.103* | -.27** | -.0552 | -.0944 | - | - | - |
| | (.0594) | (.11) | (.0567) | (.0733) | | | |
| Obs | 66263 | 13257 | 37117 | 11461 | 66263 | 66263 | 66263 |

[a] Regressions include In-school survey covariates from Table I (except where collinear by definition).

* significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors in parentheses

[c] School level covariates used in columns 1,2,3,4; grade level covariates are used in columns 5,6,7.

**Table III**
**Connectedness and Long Run Outcomes[a,b]**

| Dependent Variable | Yrs Schooling (OLS) 1 | College (Probit) 2 | Employed (Probit) 3 | Yrs Schooling (OLS) 4 | College (Probit) 5 | Employed (Probit) 6 |
|---|---|---|---|---|---|---|
| Links (ind) | .0425*** | .0342*** | .0162*** | - | - | - |
| | (.00431) | (.00367) | (.0054) | | | |
| | | [.0098] | [.0047] | | | |
| Links (sch) | | | | .215*** | .166*** | .0819** |
| | | | | (.0495) | (.0405) | (.0355) |
| | | | | | [.0483] | [.0237] |
| Parental. Education (ind) | .127*** | .11*** | .00835 | .133*** | .114*** | .0108 |
| | (.0124) | (.0104) | (.0136) | (.0128) | (.0107) | (.0137) |
| Family Income (ind) | .00196*** | .00433*** | .000642 | .00229*** | .00468*** | .000755 |
| | (.000649) | (.00113) | (.000891) | (.000696) | (.00121) | (.000904) |
| GPA(ind) | .696*** | .542*** | .0651** | .73*** | .565*** | .0733** |
| | (.0404) | (.0303) | (.0326) | (.0411) | (.0303) | (.0322) |
| AH Vocab (ind) | .0147*** | .0137*** | .00371** | .0154*** | .0142*** | .00403** |
| | (.00161) | (.00172) | (.0018) | (.00168) | (.00174) | (.00181) |
| Distance (ind) | -.000547 | -.00407 | -.00109 | -.00266 | -.00555* | -.002 |
| | (.00396) | (.00311) | (.00351) | (.00393) | (.00304) | (.00341) |
| Ave. Parent Ed. (sch) | .468*** | .419*** | .109* | .454*** | .405*** | .101* |
| | (.0615) | (.0598) | (.0575) | (.062) | (.0585) | (.0579) |
| Std. Dev. Parent Ed. (sch) | .0708 | .00429 | .171 | .0484 | -.0289 | .151 |
| | (.225) | (.182) | (.142) | (.229) | (.181) | (.142) |
| Het. Race (sch) | -.366 | -.243 | -.294* | -.213 | -.118 | -.231 |
| | (.248) | (.212) | (.174) | (.243) | (.21) | (.183) |
| Ave. GPA (sch) | -.442*** | -.493*** | -.0851 | -.528*** | -.547*** | -.117 |
| | (.156) | (.131) | (.112) | (.149) | (.128) | (.109) |
| Male(sch) | .352 | -.361 | .114 | 1.11 | .198 | .429 |
| | (.804) | (.584) | (.904) | (.825) | (.591) | (.906) |
| Size (100s - sch) | .00218 | -.00549 | -.000121 | .00479 | -.00359 | .000915 |
| | (.00726) | (.00584) | (.00459) | (.00689) | (.00561) | (.00467) |
| Income: 1000s (cnty) | -.00869 | -.013 | .0221** | -.011 | -.0154* | .0211** |
| | (.00874) | (.00821) | (.00852) | (.00865) | (.00812) | (.0087) |
| Het(cnty) | .0586 | .0436 | .116 | -.128 | -.105 | .0486 |
| | (.318) | (.337) | (.292) | (.335) | (.338) | (.286) |
| Min wage(state) | .0775 | -.122 | .0924 | -.00199 | -.185 | .0615 |
| | (.153) | (.161) | (.172) | (.164) | (.168) | (.174) |
| Obs | 10090 | 10089 | 7097 | 10090 | 10089 | 7097 |

[a]Standard errors in parentheses. Regressions include the full set of covariates from Table I.

[b]Weighted population mean of marginal effects in brackets.

* significant at 10%; ** significant at 5%; *** significant at 1%

**Table IV**
**Connectedness and Long Run Outcomes - School Fixed Effects[a,b]**

| Dependent Variable | Yrs Schooling (OLS) 1 | College (Probit) 2 | Employed (Probit) 3 | Yrs Schooling (OLS) 4 | College (Probit) 5 | Employed (Probit) 6 |
|---|---|---|---|---|---|---|
| Links (ind) | .039*** | .033*** | .014** | - | - | - |
| | (.00442) | (.00373) | (.00582) | | | |
| | | [.0092] | [.0039] | | | |
| Links (grd) | - | - | - | .126*** | .0708** | -.0132 |
| | | | | (.0288) | (.0286) | (.0308) |
| | | | | | [.0199] | [-.0036] |
| Parental. Education (ind) | .126*** | .111*** | .00824 | .133*** | .116*** | .0103 |
| | (.012) | (.0111) | (.0138) | (.0126) | (.0114) | (.0138) |
| Family Income (ind) | .00209*** | .00459*** | .000837 | .00245*** | .00495*** | .000948 |
| | (.000639) | (.0012) | (.000911) | (.000683) | (.00129) | (.000923) |
| GPA(ind) | .72*** | .574*** | .0848** | .753*** | .596*** | .0928*** |
| | (.0406) | (.0319) | (.0341) | (.0412) | (.0317) | (.0336) |
| AH Vocab (ind) | .0134*** | .0132*** | .0036* | .0142*** | .0137*** | .00383* |
| | (.00152) | (.00177) | (.00195) | (.00158) | (.00178) | (.00196) |
| Distance (ind) | -.0062* | -.00955*** | -.00117 | -.00799** | -.0107*** | -.00173 |
| | (.00369) | (.0034) | (.00432) | (.00376) | (.00345) | (.00415) |
| Ave. Parent Ed. (grd) | -.0195 | .0306 | .0344 | -.112 | -.0217 | .0434 |
| | (.121) | (.118) | (.128) | (.126) | (.119) | (.134) |
| Std. Dev. Parent Ed. (grd) | .376** | .221 | -.266 | .408** | .223 | -.28* |
| | (.182) | (.164) | (.169) | (.178) | (.162) | (.168) |
| Het. Race (grd) | .0709 | .393 | -.278 | .404 | .577 | -.323 |
| | (.799) | (.653) | (.78) | (.758) | (.661) | (.787) |
| Ave. GPA (grd) | -.365** | -.193 | -.258 | -.348** | -.189 | -.271 |
| | (.163) | (.133) | (.174) | (.157) | (.131) | (.173) |
| Male(grd) | -.0638 | .00896 | -.0368 | .0706 | .0579 | -.0573 |
| | (.434) | (.348) | (.356) | (.406) | (.331) | (.367) |
| Size (100s - grd) | -.000467 | -.00123 | .000663 | -.000921 | -.00143* | .000718 |
| | (.000896) | (.000778) | (.000587) | (.000859) | (.000728) | (.000614) |
| Obs | 10090 | 10090 | 7097 | 10090 | 10090 | 7097 |

[a]Standard errors in parentheses. Regressions include full set of covariates from Table I (except where collinear by definition).

[b]Weighted population mean of marginal effects in brackets.

* significant at 10%; ** significant at 5%; *** significant at 1%

**Table V**

**Connections, GPA, and Education Attainment**[a]

| Dependent Variable: | GPA (OLS) 1 | Grad HS (Probit) 2 | College (Probit) 2 |
|---|---|---|---|
| **A. Ave. GPA (peers) < Ave. GPA (School)** | | | |
| Links (ind) | -.00449 | .0169*** | .0152** |
| | (.00292) | (.00591) | (.00684) |
| Ave. GPA (peers) | .361*** | .32** | .532*** |
| | (.0602) | (.14) | (.109) |
| Obs | 4268 | 4268 | 4268 |
| **B. Ave. GPA (peers) > Ave. GPA (School)** | | | |
| Links (ind) | .0188*** | .0471*** | .0328*** |
| | (.00245) | (.00525) | (.00854) |
| Ave. GPA (peers) | .59*** | .353*** | .33** |
| | (.0655) | (.0996) | (.15) |
| Obs | 4350 | 4350 | 4350 |

[a]Standard errors in parentheses. Regressions include the full set of covariates from Table I.

* significant at 10%; ** significant at 5%; *** significant at 1%

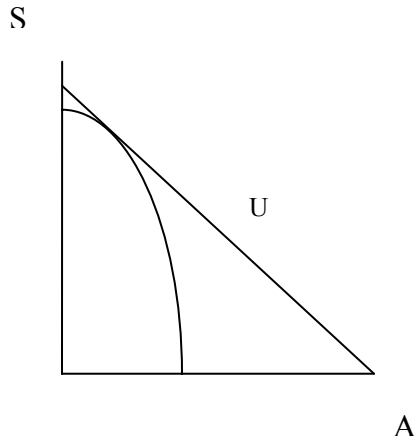
**Table VI**
**Connectedness and School Policies**

| | Links (Ave) | Count |
|---|---|---|
| Magnet schools | 6.83<br>(2.04) | 15 |
| Schools with Busing to Achieve Racial Targets | 6.23<br>(1.49) | 6 |
| All Other Schools | 8.87<br>(1.90) | 108 |

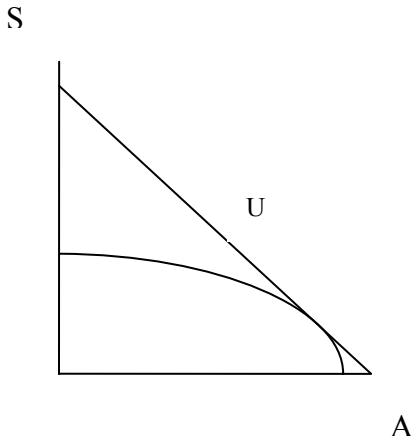
Standard Deviation in parentheses

**Figure I**

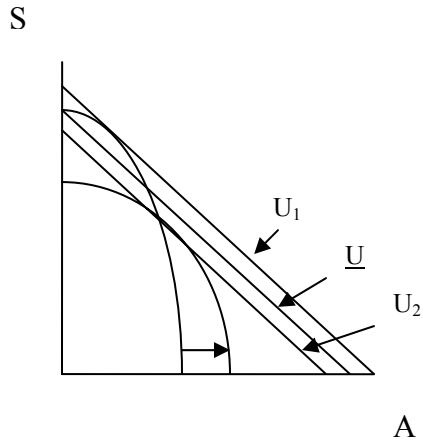
A. $\theta_s > \theta_a$



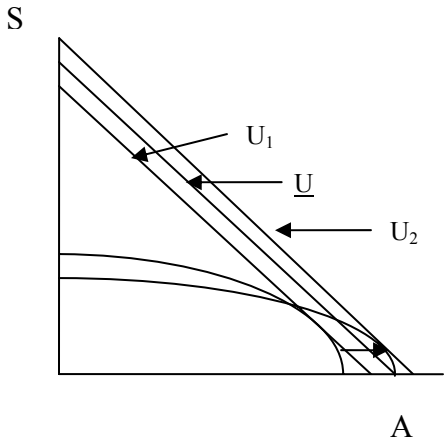

B. $\theta_a > \theta_s$



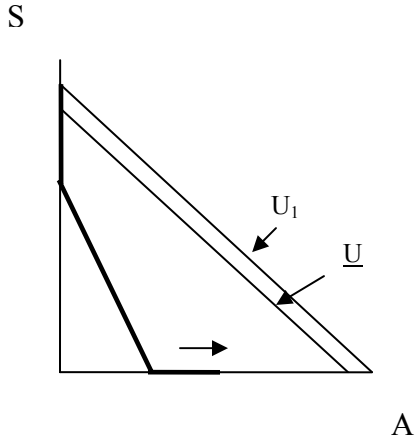

C.




D.




E.




F.

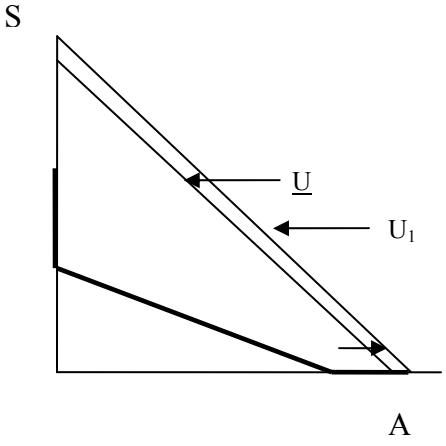
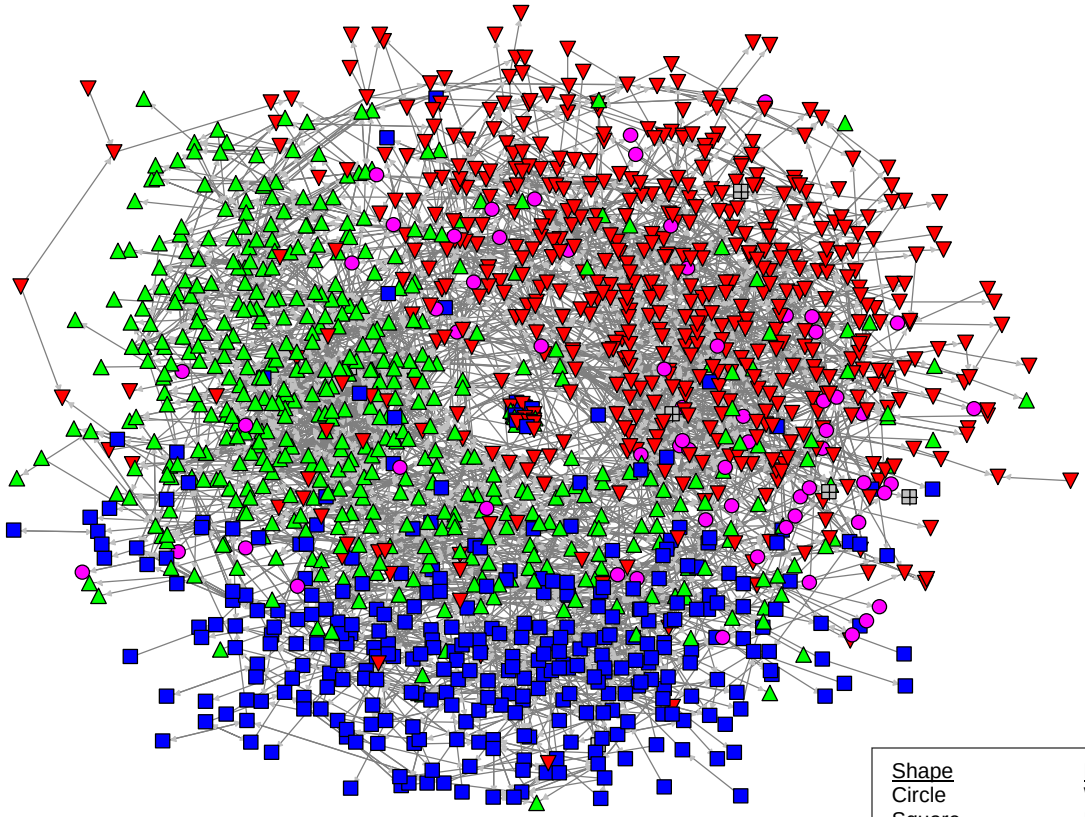


**Figure II**

**A. Sch 77: Het=.68 Links=5.9**



| Shape | Race |
|---|---|
| Circle | White |
| Square | Black |
| Up Triangle | Asian |
| Dwn Triangle | Hispanic |
| Box | Amer. Indian |

**B. School 146 : Het=.37 Links=8.74**

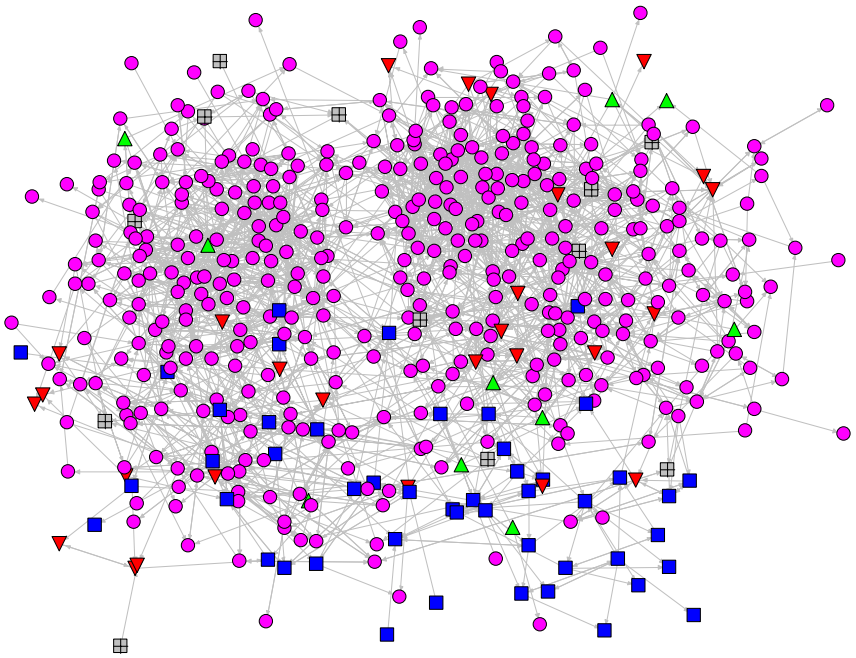


**Figure III**

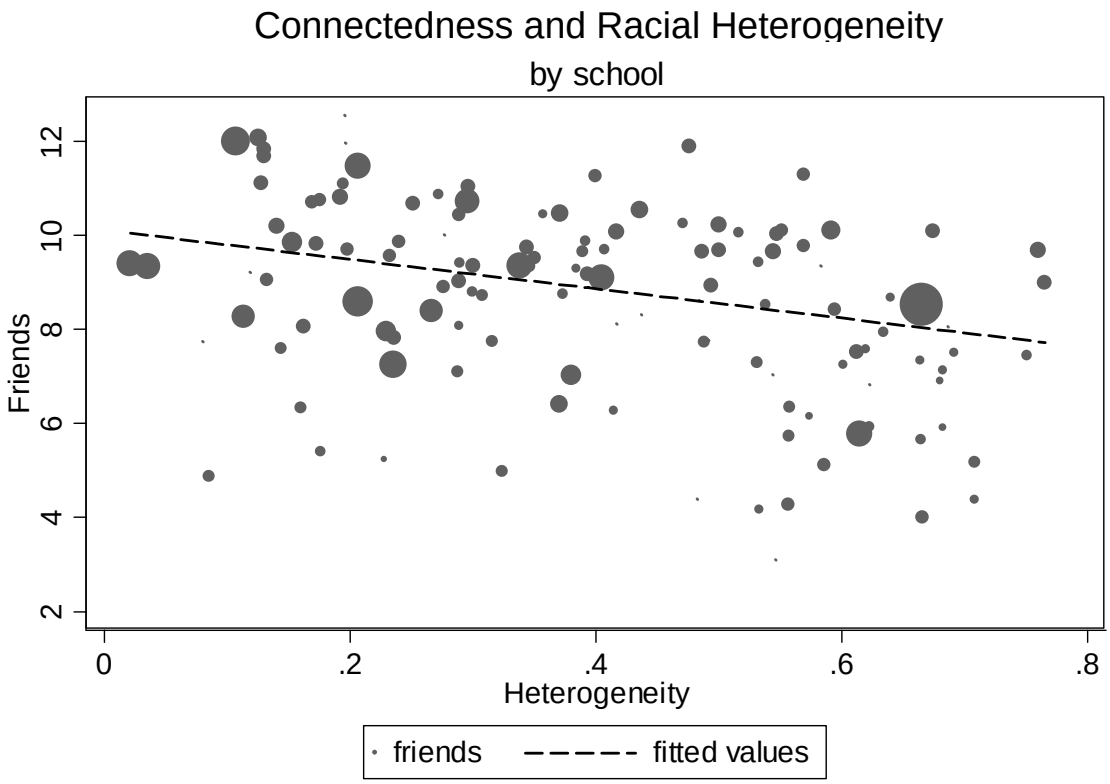
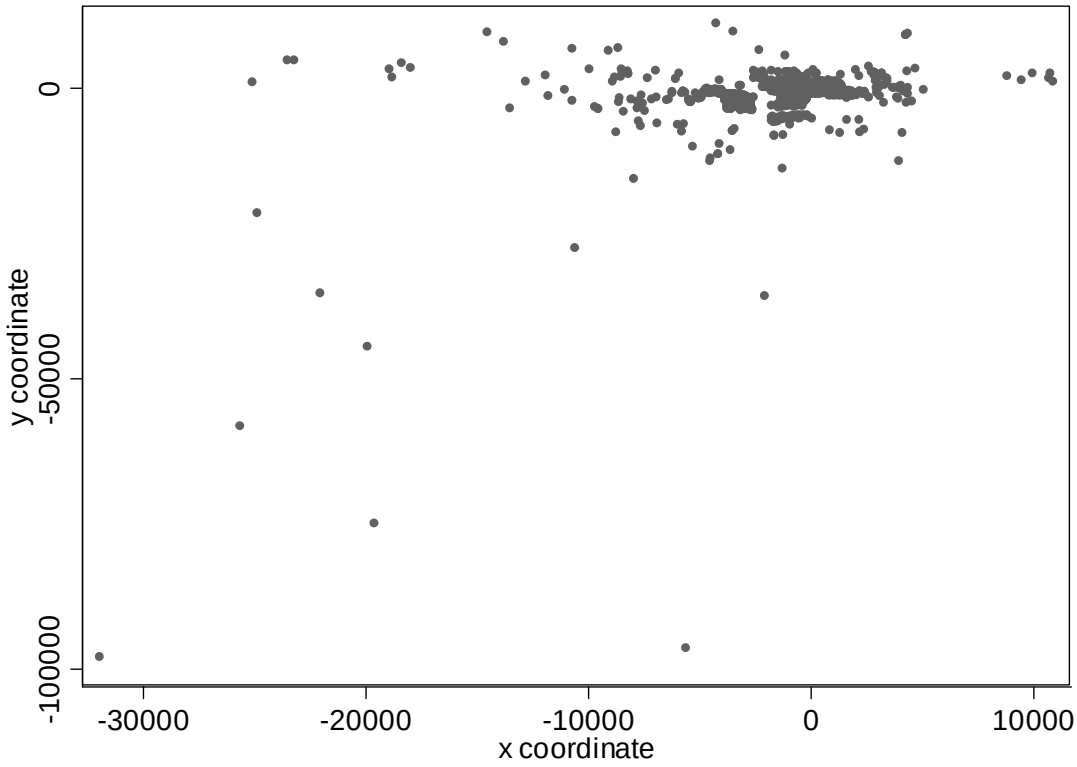

Connectedness and Racial Heterogeneity
by school
Friends
2
4
6
8
10
12
0
.2
.4
.6
.8
Heterogeneity
friends
fitted values

**Figure IV**
**School Residences***




*x, y coordinates in meters