Friend Effects and Racial Disparities in Academic Achievement
Jennifer Flashman
University of Notre Dame

Abstract: Racial disparities in achievement are a persistent fact of the US educational system. An often cited but rarely directly studied explanation for these disparities is that adolescents from different racial and ethnic backgrounds are exposed to different peers and have different friends. In this article I identify the impact of friends on racial and ethnic achievement disparities. Using data from Add Health and an instrumental variable approach, I show that the achievement characteristics of youths’ friends drive friend effects; adolescents with friends with higher grades are more likely to increase their grades compared to those with lower-achieving friends. Although these effects do not differ across race/ethnicity, given differences in friendship patterns, if black and Latino adolescents had friends with the achievement characteristics of white students, the GPA gap would be 17 to 19 percent smaller. Although modest, this effect represents an important and often overlooked source of difference among black and Latino youth.

Keywords: friendship; friend effects; achievement gaps; race/ethnicity; academic achievement; instrumental variables

Background

Minority students are increasingly isolated in racially and socioeconomically segregated schools (Orfield, Kucsera, and Genevieve 2012; Orfield and Lee 2007). Today, more than 80 percent of black students and 74 percent of Latino students attend schools that are majority minority, and more than 60 percent of their classmates come from low-income households. Recent research connects these patterns of segregation to racial and ethnic achievement gaps (Berends and Penaloza 2010; Brown-Jeffy 2006; Card and Rothstein 2007), arguing that youths’ peer environments and friendship opportunities are key mechanisms linking segregation and achievement (Bifulco and Ladd 2006; Grissmer, Flanagan, and Williamson 1998).

Friends play an important role in children’s and adolescents’ lives, providing support, resources, and information (Coleman 1988; Portes 1998). While theoretically these friendships have the potential to decrease the distance between the most and least advantaged, segregation restricts the opportunities that individuals have for friendships (Blau 1977; Quillian and Campbell 2003) and leads to systematic variation in the characteristics of adolescents’ friends by racial and ethnicity (Crosnoe, Cavanagh, and Elder 2003; Flashman 2012; Fryer and Torelli 2010). Yet, in spite of the theoretical importance of friends, their effect on achievement gaps is rarely directly studied.

This article has two goals: 1) to provide a causal estimate of the effect that friends have on academic achievement and 2) to connect these causal effects to patterns of racial and ethnic inequality in academic achievement. Using data from the National Longitudinal Study of Adolescent Health and instrumental variable analysis to deal with issues of endogeneity and selection, I identify the effect of friends’ characteristics on grades, paying particular attention to differences in friend effects across race/ethnicity. I use these estimated friend effects and differences in the characteristics of black and Latino adolescents’ friends to explain GPA differences across race/ethnicity, providing an important look at how the social settings of schools matter for racial and ethnic stratification.

Background

The Foundations of Friend Influence

Friends affect the academic behaviors and outcomes of youth through two key pathways. First,
friends create and reinforce common norms and expectations of behavior (Coleman 1988). Through a combination of role modeling and sanctions, youth develop norms affecting educational behaviors and achievement (Sewell, Haller, and Portes 1969; Steinberg, Brown, and Dornbusch 1996). Second, friendships provide channels for the diffusion of information (Coleman 1988; Granovetter 1973; Portes 1998). By helping each other with homework or sharing information about college, courses, or teachers, friends can help to improve academic outcomes and change educational behaviors (Crosnoe et al. 2003).

Both mechanisms are invoked in recent research studying friends and their role in the academic achievement and educational attainment of youth. This work shows that friends have important and large associations with adolescents’ academic outcomes. Academically inclined friends insulate youth from failure (Crosnoe et al. 2003), increase the probability of enrolling in advanced courses and high-level math (Crosnoe et al. 2008; Frank et al. 2008; Riegle-Crumb, Farkas, and Muller 2006), and are particularly beneficial for low-achieving students (Altermatt and Pomerantz 2005). Overall, having friends who do well in school is a strong predictor of grades (Lin 2010). Similarly, having friends or peers with high levels of parental education increases the likelihood of enrolling in college (Choi et al. 2008) and has a positive impact on grades in school (Fletcher, Ross, and Zhang 2013; Riegle-Crumb and Callahan 2009). Through both the creation of norms and the sharing of information, the academic and socioeconomic characteristics of adolescents’ friends are associated with achievement trajectories.

Race, Ethnicity, and Friendship

Theoretically, friends are especially important for racial and ethnic minorities (Carter 2005; Crosnoe et al. 2003; Riegle-Crumb and Callahan 2009; Stanton-Salazar and Dornbusch 1995). On the one hand, the characteristics of friends differ systematically by race/ethnicity; as a result, friendships may inhibit the educational trajectories of black and Latino youth relative to whites. Both the grades and parental education of black and Latino adolescents’ friends are lower than those of white and Asians (Crosnoe et al. 2003; Flashman 2012; Fryer and Torelli 2010). If, as argued above, the academic and socioeconomic characteristics of friends impact academic performance and educational attainment, then the composition of black and Latino youths’ friends will depress their grades and future attainment.

Moreover, black and Latino adolescents are significantly more likely to be friends with other co-ethnics (Moody 2001; Quillian and Campbell 2003). Ogbu (1994) argues that co-ethnic relationships among black and, to a lesser extent, Latino students support and encourage an oppositional culture where academic achievement and attainment are discouraged. Having more black or Latino friends will strengthen this oppositional culture and lead to lower grades and more negative schooling behaviors.

While the compositions of students’ friends differ by race/ethnicity, heterogeneity in the effect of friends may compound this disadvantage. The structural features of adolescents’ networks differ across race/ethnicity, and those differences affect the ways in which friends are able to impact grades and college-going. In the data used in this article, for example, black adolescents’ friendships are reciprocated less often, black students named fewer in-school friends on the survey, and black students are less central in their in-school networks. These differences impact the extent to which friends can influence behaviors by affecting information flows (Calvo-Armengol, Patacchini, and Zenou 2009; Moody and White 2003), suggesting that black and Latino students’ friendships will exert less influence on their academic performance and educational attainment. Supporting this argument (though not testing the mechanism), Alvarado and Turley (2012) show that the association between friends and college application is smaller for Latino youth than for white youth.

On the other hand, in spite of their friends’ academic and socioeconomic disadvantages, friends may be more beneficial to black and Latino youth. The immigrant experience of many Latino youth, for example, may be a factor. First- and second-generation immigrants have strong ties with extended family networks and co-ethnic communities (Portes and Rumbaut 2001). These networks provide important resources and information for migration decisions (Palloni et al. 2001), navigation of the labor market (Nee and
Sanders 2001), and educational decisions (Kastenutz et al. 2008). Children in immigrant communities, observing the central role of networks, may similarly turn to their own networks for support, information, and resources. Riegle-Crumb and Callahan (2009), indeed, show that Latino students with more co-ethnic friends have more positive achievement outcomes. Moreover, Crossnoe et al. (2003) argue that while friends provide a redundant resource for white students, whose parents have more resources and information to draw upon, friends are often the only source of information and support available to black students. Because black students are so dependent on their friends, their friends should have stronger impacts on their academic behaviors and grades.

These arguments suggest that friends may influence racial and ethnic differences in achievement, and grades in particular, in two ways: 1) through compositional differences in friends’ grades, socioeconomic status, and race/ethnicity and 2) through differences in the effects of friends’ grades, socioeconomic status, and race/ethnicity. The direction and importance of each element, however, are unclear.

Methodological Challenges

Complicating matters, studying friends comes with enormous methodological challenges. Individuals choose their friends and tend to choose friends with similar attributes to themselves (selection) (McPherson, Smith-Lovin, and Cook 2001), and these choices affect the behaviors of all participants in the friendship (endogeneity) (Manski 2000). As a result, although I may observe that two friends are similar with respect to grades, I cannot know the degree to which this pattern is the consequence of the influence of friend 1 on friend 2, friend 2 on friend 1, or selection based on similarity. Although the majority of sociological research on friends uses panel data, little is done to address these serious selection and endogeneity issues.

A number of recent economics articles begin to address these methodological issues by using features of friend and peer networks to address endogeneity (Bamoulle, Djebbari, and Fortin 2009; Lin 2010) and selection (Fletcher et al. 2013). Fletcher et al. (2013) take a promising approach using instrumental variables to identify friend effects and show that female students whose friends’ mothers have high levels of education have higher self-reported grades.1

Although this work provides an important first step in identifying the causal effect of friends, it focuses on friend effects by studying the educational composition of their parents. As Fletcher et al. note, the likely mechanism for this effect is through those friends’ educational behaviors and academic achievements. If we want to understand not just whether but which characteristics of friends matter, it is vital that we identify the extent to which friend effects are driven by friends’ grades and if other compositional characteristics of adolescents’ friends, such as their socioeconomic background, have independent effects. Understanding these relationships in a causal framework is key to identifying the contribution of friends to racial and ethnic differences in achievement generally and grades in particular.

Research Questions

While research consistently shows that the academic, socioeconomic, and racial and ethnic characteristics of friends are associated with the educational behaviors of youth, it remains unclear whether and how these patterns contribute to achievement gaps across race/ethnicity. In light of these limitations, I ask three key research questions:

1. How do the characteristics of friends influence grades?

2. Do these friend effects differ across race/ethnicity?

3. To what extent do differences in the characteristics of black and Latino adolescents’ friends and their effects explain GPA differences across race/ethnicity?

I address these questions, paying particular attention to the methodological challenges I face by studying friends. I use a combination of instrumental variables and lagged measures of

1 The authors follow a two-step process to construct an instrument. They first use school-by-grade variations in the compositions of students’ cohort-mates to predict the characteristics of their friends. They then use the predicted number of friends whose mothers have high levels of education as an instrument for the actual number.
friend characteristics to provide causal estimates of the influence of friends’ grades, parental education, and racial and ethnic composition on future grades, accounting for both selection and endogeneity. By estimating the causal effect of friends on grades and then linking these to patterns of inequality, this article contributes to the growing literature studying friends as an important resource for children and adolescents and is the first to directly study friends as a causal mechanism contributing to achievement gaps.

Data and Measures

Data for this analysis come from the National Longitudinal Study of Adolescent Health (Add Health). Add Health provides the only nationally representative survey of adolescents and their in-school friends. In the 1994-95 school year, all 7-12th grade students in 144 sampled schools were surveyed (N=89,940). Students completed a short survey including basic socioeconomic, demographic, and achievement information. Additionally, students nominated and were linked to their five closest male friends and five closest female friends. Because nearly all students at the sampled schools participated, Add Health provides a unique dataset from which to characterize the academic composition of students’ in-school friends. A randomly selected subsample of students was followed up with in the wave 1 in-home survey in 1995, an average of six months after the in-school survey (N=20,745), and again in the spring of 1996 (wave 2, N=14,738). Twelfth graders who were surveyed in wave 1 were not included in wave 2. The follow-up waves provide longitudinal achievement data (Bearman, Jones, and Udry 1997; Chantala 2006).

Dependent Variable

To capture academic achievement and achievement differentials across race/ethnicity, I use self-reported grades in English, math, science, and social studies, recorded in the wave 2 in-home survey and transformed into grade point averages (GPA). Although transcript grades are available for a subsample of respondents, in order to maximize the sample size, I focus on self-reported GPA, which is highly correlated with transcript GPA ($r=.75$).

Key Independent Variables

The characteristics of respondents’ friends serve as the key independent variables predicting GPA. The friendship nominations from the wave 1 in-school survey are used to identify friendships. If person A nominates person B and B attends the same school as A, then B is defined as A’s friend, regardless of whether B nominates A and regardless of whether A and B are the same gender. I focus on four characteristics of friends that have been argued to impact academic performance: friends’ average GPA, average parental education, proportion black, and proportion Latino. For friends’ average GPA, friends’ self-reported grades in English, math, science, and social studies from the wave 1 in-school survey are used to construct GPAs. Those GPAs are averaged across all friends. Average parental education and race/ethnic composition are also based on friends’ self-reports of their parents’ highest level of education and their race/ethnic backgrounds, respectively.

Sample

The analysis includes all 9-11th graders from 1994-95 participating in the wave 1 in-school survey who also participated in the wave 2 in-home survey (N=7,016). Respondents who nominate zero friends’ average GPA, average parental education, proportion black, and proportion Latino.

---

2 Respondents report their average grade (A, B, C, or D, or lower) in each subject in the most recent grading period. Letter grades are assigned 4, 3, 2, and 1 points respectively and averaged across subjects.

3 Nonetheless, I replicate the analysis predicting transcript grades. Results from models predicting transcript GPA from 1995-96 are included in Table A.8 of the online supplement.

4 Including all friend nominations rather than only reciprocated friend nominations addresses some of the error in friend reports that results from survey constraints on the number of nominations allowed.

5 I investigated sensitivity to this measurement decision, replicating the analysis using the proportion of friends with GPAs greater than 3.0 and the proportion of friends with a college-educated parent. Results are substantively consistent across these alternative measures and can be found in Table A.9 of the online supplement.

6 A total of 10,703 respondents participated in the wave 1 in-school and wave 2 in-home surveys. Of those, 3,675 are in 7th, 8th, or 12th grade in the 1994-95 school year. Because the majority of 7th and 8th graders are in separate middle schools where the stakes for good grades are lower, the standards for those grades are different than
in-school friends are excluded from the analysis (N=1,296), as are those who are missing self-reported grades (W1: N=501, W2: N=340) and those whose friends and indirect friends are missing information on any of the key friend characteristics or instruments (described below) (N=141 and N=87 respectively). A summary of these exclusions is included in Table A.2 in the online supplement. With these exclusions the final analysis sample includes 4,651 respondents in 101 schools.

Methods

Equation 1 describes the GPA, $Y_i$, of individual $i$ with friends $n$ as a function of individual characteristics $X$, friend characteristics $A$, and error term $\varepsilon$, or

$$Y_{in} = \beta_1 X_i + \beta_2 A_{in} + \varepsilon_i$$  \hspace{1cm} (1)

As I mention earlier, two problems occur when estimating Equation 1. First, individuals choose their friends and tend to choose friends with similar characteristics to themselves (McPherson et al. 2001). If I do not fully account for the factors that lead to friend choices, then I will incorrectly attribute the similarity observed between two individuals on a behavior to influence, and overestimate the effect of friends on achievement (Mouw 2006).

Second, Equation 1 imposes a causal order. The friend characteristics, $A_{in}$, affect individual $i$’s achievement, $Y_{in}$. However, if the outcome of each friend is also governed by Equation 1, then $A_{in}$ would also be determined by $Y_{in}$ through individual $i$’s influence, or

$$A_{in} = \beta_1 X_i + \beta_2 Y_{in} + \varepsilon_i$$  \hspace{1cm} (2)

If $Y_{in}$ and $A_{in}$ occur simultaneously, the causal order goes in both directions at once. As a result, adolescents will appear more similar, and that similarity will be attributed to the influence of $A_{in}$, upwardly biasing estimates of the effect of friend characteristics, $\beta_2$ (Manski 2000).

Studying the socioeconomic characteristics of friends’ parents—as Fletcher et al. (2013) do—ensures that this endogeneity is not a problem; an individual will not influence his friends’ parents’ level of education. However, in order to get at the potential mechanism of the friend effect—in this instance the youths’ friends’ academic performance—endogeneity is introduced.

Estimation Strategy

In order to address endogeneity, the characteristics of friends as reported in the 1994-95 school year are used to predict GPA reported in the 1995-96 school year. Studying the change in GPA that results from friend characteristics at an earlier point in time ensures that the causal direction of the friend effect goes from friends’ behaviors to the respondent’s and not the other way around (Manski 2000; Mouw 2006). This lag may lead to an underestimate of friend effects by leaving out concurrent friend effects (Hanushek et al. 2003). As a result, these estimates can be understood as conservative estimates of friend effects.

In addition, I use a series of instrumental variables to deal with selection. For Equation 1, I need to identify a variable, $Z_{in}$, that predicts $A_{in}$ but is uncorrelated with $\varepsilon_i$ (Bollen 2012; Murnane and Willett 2011). I use the characteristics of an individual’s friends’ friends (indirect friends) who are not the focal individual’s friends as an instrument for the four characteristics of adolescents’ friends identified earlier. Figure 1 illustrates the relationship between an individual, her friends, and her indirect friends in a hypothetical network. For individual A, individuals B, C, D, and E are her friends (the grey circles). Several of those friends are friends with individuals who are not friends with A. These friends of friends (the black circles)—F, G, H, I, and J—are A’s indirect friends. The average characteristics of these indirect friends are used as instruments for the average characteristics of A’s direct friends.

The number of indirect friends of respondents ranges from one to 96, and the average respondent has 30 indirect friends (standard deviation=19). Self-reported achievement, parental education, and expectations of graduating from college are averaged across all indirect friends, and the proportions of indirect friends from two-parent families, who are black, and who are Latino are calculated. Together, these indirect friend characteristics are used to instrument the average self-reported GPA, parental education, and racial
composition of respondents’ nominated friends at the time of the wave 1 in-school survey. Table 1 summarizes the GPA, friend characteristics, and indirect friend characteristics of respondents overall and by race/ethnicity.

Two general assumptions must hold for any instrument to be a valid, in this case: 1) the indirect friends’ characteristics must be independent of the error term when predicting grades and 2) one’s friends’ characteristics must be correlated with her indirect friends’ characteristics. The second assumption is relatively easy to establish; the first is more challenging. Additionally, several smaller issues unique to this set of instruments must be addressed, specifically error in the measurement of friendships and issues of transitivity and friend dynamics. I discuss the two general assumptions before returning to the issues unique to this case.

Assumption 1: The instruments are independent of the error term.

For the indirect friends to serve as an instrument for friend characteristics, their characteristics must be unrelated to $\varepsilon$, the error term. While this assumption cannot be tested directly, with a rich set of controls and arguments for how these controls account for the ways that indirect friends’ characteristics are correlated with respondents’ future grades, this instrument can provide a novel way to estimate the causal effect of friends.

In the absence of controls, the indirect friend characteristics are surely correlated with $\varepsilon$. Characteristics of individuals’ families, their course-taking, and their educational involvement, among other things, will all affect both who their indirect friends are and their future GPAs, even after controlling for friend characteristics. To account for these relationships I include three groups of controls in the analysis: demographic and family background characteristics, course-taking and school involvement, and attitudes and expectations towards school. In addition, self-reported GPA at the time of the in-school survey is included to control for prior achievement, the role that friends may have played in affecting those grades, the role that the respondents’ grades may have played in affecting their friends’ grades, and the role of achievement in affecting friendship choices.

No instrument is perfect, and a major concern about using indirect friends as an instrument for friends is the implicit assumption that the classmates of respondents will have no direct effect on respondents’ grades, except through the characteristics of their friends (and other controls). The extensive research on peer effects suggests that classmates play an important role in the

---

7 A description of the control variables is included in Table A.1 of the online supplement. All controls are measured at the time of the wave 1 in-school survey, when friend nominations occur. As a result, a few key controls are not included in this analysis: specifically, parents’ educational expectations, self-reported intelligence, and scores on the Add Health picture vocabulary test. Because they are measured after friend nominations occur, these measures are potentially endogenous to the friend effects.
course-taking patterns and test scores of adolescents (Frank et al. 2008; Hanushek et al. 2003; Hoxby 2000). If classmates play a similarly important and independent role in affecting GPA, then this set of instruments, even with the controls described above, will be invalid. In order to account for these broader peer effects, I additionally include school-by-grade fixed effects. All measured and unmeasured characteristics of adolescents’ schools and classes, shared by their grade-mates, are controlled. If the characteristics of adolescents’ broader peer environments matter independently of their friends, these fixed effects will account for these influences.

Still, to play devil’s advocate, I can imagine several ways in which the characteristics of indirect friends may nevertheless be correlated with the error term. First, a student’s indirect friends represent his or her possible future friends. In order to facilitate those future friendships, youth may emulate the behaviors of their indirect friends, influencing their educational behaviors and future GPA. Indeed, Christakis and colleagues show that the smoking behavior, levels of depression, and obesity of individuals and their indirect friends are correlated (Cacioppo, Fowler, and Christakis 2009; Christakis and Fowler 2007; Christakis and Fowler 2008). A similar correlation may exist among youth and their indirect friends’ academic behaviors. However, for this argument to undermine the instrument, this effect needs to flow directly from the indirect friend to the respondent, independent of their friend characteristics and other controls. Because the indirect friend is connected to the respondent through their direct friend, and their direct friends’ attitudes and behaviors resemble those of the indirect friend, much of this effect likely filters through the direct friend or through contextual effects. Supporting this argument, Cohen-Cole and Fletcher (2008a; 2008b) show that the network effects claimed by Christakis and Fowler to affect obesity disappear after controlling for context and selection.

Second, the respondent may directly influence friends’ friend choices, thereby introducing a correlation between indirect friends and the error term. For example, youth may affect their friends’ friend choices by stigmatizing or popularizing their friends. If the respondent is perceived by others as unpopular, his or her friends may be stigmatized by association, affecting the friend

### Table 1: Means and standard deviations of individual GPA, friends’ characteristics, and indirect friends’ characteristics, Add Health

<table>
<thead>
<tr>
<th></th>
<th>Full Sample</th>
<th>White</th>
<th>Black</th>
<th>Latino</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>S.D.</td>
<td>Mean</td>
<td>S.D.</td>
</tr>
<tr>
<td>Self-Reported GPA Spring</td>
<td>2.87</td>
<td>0.73</td>
<td>2.93</td>
<td>0.74</td>
</tr>
<tr>
<td>Friend Characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average GPA W1</td>
<td>2.83</td>
<td>0.52</td>
<td>2.89</td>
<td>0.51</td>
</tr>
<tr>
<td>Average Parental Education</td>
<td>3.11</td>
<td>0.85</td>
<td>3.18</td>
<td>0.77</td>
</tr>
<tr>
<td>Proportion Latino</td>
<td>0.16</td>
<td>0.29</td>
<td>0.06</td>
<td>0.13</td>
</tr>
<tr>
<td>Proportion Black</td>
<td>0.17</td>
<td>0.34</td>
<td>0.02</td>
<td>0.09</td>
</tr>
<tr>
<td>Indirect Friend Characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average GPA W1</td>
<td>2.85</td>
<td>0.41</td>
<td>2.92</td>
<td>0.38</td>
</tr>
<tr>
<td>Proportion Two Parent Families</td>
<td>0.75</td>
<td>0.18</td>
<td>0.81</td>
<td>0.13</td>
</tr>
<tr>
<td>Average Coll. Grad. Expectations</td>
<td>6.58</td>
<td>0.94</td>
<td>6.55</td>
<td>0.87</td>
</tr>
<tr>
<td>Average Parental Education</td>
<td>3.13</td>
<td>0.68</td>
<td>3.20</td>
<td>0.60</td>
</tr>
<tr>
<td>Proportion Latino</td>
<td>0.16</td>
<td>0.27</td>
<td>0.06</td>
<td>0.10</td>
</tr>
<tr>
<td>Proportion Black</td>
<td>0.17</td>
<td>0.32</td>
<td>0.03</td>
<td>0.07</td>
</tr>
<tr>
<td>N</td>
<td>4,651</td>
<td>2,651</td>
<td>854</td>
<td>688</td>
</tr>
</tbody>
</table>

*Note:* The full sample includes all 9-11th grade students who participated in the W1 In School and W2 In-home surveys, nominated at least 1 in-school friend, and are not missing self-reported GPA.
options available to the stigmatized friend. That said, many of the characteristics that stigmatize (or popularize) students are measured and controlled in the model, including prior achievement, course-taking, club membership, participation in sports, attachment to school, etc. (Crosnoe 2011). To the extent that these characteristics capture the key ways that the respondent’s reputation affects friends’ friend choices, this issue is dealt with. Similarly, adolescents may directly affect their friends’ friend choices by affecting their activities and course-taking patterns, and thus the group of people to which their friends are exposed. Again, including controls to account for the respondents’ club membership, course-taking, and sports teams helps account for this form of influence over indirect friends.

Although the assumption of independence between the instrument and the error term cannot be tested, the controls and school-by-grade fixed effects account for many of the ways that respondents affect and are affected by indirect friends. Any additional correlation between indirect friends and the error term is assumed to be negligible.

Assumption 2: The set of instruments provide a good prediction of the endogenous variables.

The second assumption states that the characteristics of indirect friends must be correlated with the characteristics of respondents’ chosen friends. Unlike the first assumption, this can be tested directly. Based on the F-statistic shown at the bottom of Table 2, the characteristics of respondents’ indirect friends are strongly correlated with the endogenous friend characteristics, even after the set of controls and fixed effects are included.

Measurement Error

As I note above, several additional issues specific to the case of indirect friends must be addressed for these instruments to be valid. First, Add Health places a limit on the number of friends that respondents are allowed to nominate. A consequence of these constraints is that given the ability to nominate more friends, the respondent might in fact have chosen at least some of the indirect friends as friends, with the result that those indirect friends’ characteristics are subject to the same selection bias as the nominated friends.

Recall that the indirect friend characteristics are averaged and are averaged across a large number of individuals. To the extent that nomination error is only a small problem, the characteristics of the actual friends who are being treated as indirect friends are averaged with the rest of the indirect friends, minimizing the impact.

Transitivity and Friend Dynamics

A well established pattern in social networks research is the tendency, known as transitivity, for the friends of an individual’s friends to become friends with that individual (Goodreau, Kitts, and Morris 2009). Between the time that respondents’ friendships are observed and the time that their future grades are observed, more than a year later, indirect friends could become direct friends, influencing the respondents’ behaviors. Although the majority of the friendship data in Add Health is cross-sectional, longitudinal friendship data were collected in a saturation sample of 16 schools. In these schools I assess the probability that an individual becomes friends with her indirect friends between waves 1 and 2. Approximately 5 percent of respondents in the saturation sample had at least one friend at a later wave who was a friend of a friend at the time of the wave 1 survey. These descriptive results are mirrored in a network analysis that shows that the predicted probability of creating a new friendship with an individual who is a friend of a friend is 0.58 and

9 A more liberal definition of friendship can minimize measurement error resulting from nomination limits. The current definition allows for error in respondents’ friends’ reports by not requiring reciprocation. The friendship definition can be further relaxed by including those individuals who nominate the respondent as a friend but are not nominated by the respondent. Although this definition cannot account for all error in friend reporting, and likely introduces some error (creating friendships where there are none), if individuals prefer reciprocated friendships, (Rivera, Soderstrom, and Uzzi 2010) then these un reciprocated nominations of the respondent are a likely set of actual friendships not reported by respondents because of limitations on the number of nominations allowed. I replicate the analysis with this more expansive definition to ensure that the substance of my results is not sensitive to this form of measurement error. These results are presented in Table A.11 in the online supplement.
the probability of creating a new tie with an individual who is not a friend of a friend is 0.42. Although transitivity does matter to the creation of friendships, it does not overwhelm friendship choices, and in practice it leads to only a small number of realized friendships. These patterns suggest that, while transitivity should be a concern with these instruments, it ultimately will not cause issues if the analysis sample follows the pattern of the saturation sample.

The instrumental variable analysis described above is implemented using two-stage least squares regression in STATA 13. All six of the indirect friend characteristics are used to estimate each of the endogenous friend characteristics. Standard errors are clustered at the school level. Missing data flags are included in the analysis for the few respondents missing data on controls. In all cases, those missing data on a particular control do not differ significantly in their achievement from those not missing data. Table A.3 in the online supplement includes an analysis of the differences between the analytic sample (described earlier) and the respondents excluded from the analysis. In general, the analytic sample is more advantaged. They come from more educated families, are higher achieving, are more involved in school, and take more advanced math and English courses. Given these features of the sample, I am likely underrepresenting the role that friends play in affecting inequality. The most disadvantaged minority students are underrepresented in the sample, thereby reducing the GPA gaps across race and ethnicity. If friends contribute to achievement gaps, the fact that the bottom of the distribution is excluded from the analysis will lead friends to appear less important than they in fact are.10

Results

The Effect of Average Friendship Group GPA

The first two columns of Table 2 present the results of two OLS regressions, showing the simple relationship between friend characteristics at wave 1 and GPA at wave 2. In model 1, which does not include controls or fixed effects, both the grades of students’ friends and their parental education are associated with respondents’ future grades. A positive one point difference in friends’ average GPA at wave 1 is associated with a 0.47 point increase in a respondent’s future GPA, and an individual with a group of friends whose average parental education is, for instance, some college versus a high school degree, has a GPA 0.07 grade points higher on average. In both cases, these relationships are statistically significant at the p < 0.05 level. The proportion of black friends and the proportion of Latino friends are not significantly related to grades in the following year.

Model 2 adds controls and school-by-grade fixed effects. With the addition of these controls the association between friends’ GPA and GPA a year later weakens, and the association between friends’ parents’ education and future GPA becomes insignificant. Nonetheless, even accounting for family background, course-taking, educational involvement, and all common characteristics of schools and peers experienced by classmates, the average GPA of respondents’ friends has a significant positive association with GPA a year later. A one point increase in friends’ average GPA is associated with a 0.24 grade point increase in a respondent’s future GPA.

While these estimates mitigate endogeneity to some degree—because friend characteristics are used to predict GPA in the following year—they are likely upwardly biased as a result of unmeasured selection. Model 3 shows the results of the second stage regression predicting GPA using indirect friends’ self-reported GPA, parental education, educational expectations, proportion from two-parent families, and racial composition as instruments for each of the four endogenous friend characteristics.11 These estimates represent the
Table 2: OLS and IV regression estimates predicting self-reported GPA, Add Health

<table>
<thead>
<tr>
<th>Friend Characteristics</th>
<th>(1) OLS</th>
<th>(2) OLS</th>
<th>(3) 2SLS IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average GPA W1</td>
<td>0.467‡</td>
<td>0.235‡</td>
<td>0.158‡</td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
<td>(0.027)</td>
<td>(0.043)</td>
</tr>
<tr>
<td>Average Parental Education</td>
<td>0.066†</td>
<td>0.008</td>
<td>−0.038</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.017)</td>
<td>(0.044)</td>
</tr>
<tr>
<td>Proportion Latino</td>
<td>−0.116*</td>
<td>0.023</td>
<td>−0.065</td>
</tr>
<tr>
<td></td>
<td>(0.066)</td>
<td>(0.055)</td>
<td>(0.103)</td>
</tr>
<tr>
<td>Proportion Black</td>
<td>−0.073</td>
<td>−0.069</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>(0.056)</td>
<td>(0.076)</td>
<td>(0.084)</td>
</tr>
<tr>
<td>Controls</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>School-Grade Fixed Effects</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>F-Statistic</td>
<td></td>
<td></td>
<td>24.65</td>
</tr>
</tbody>
</table>

* p < .1; † p < .05; ‡ p < .01.

Note: N = 4,651 Standard errors in parentheses below unstandardized coefficients. Models account for clustering at the school level. Full results are available in Tables A.4 and A.5 in the online supplement.

average causal response across the population of students with friends (Angrist and Pischke 2009). The effect of friends likely varies across individuals in a multitude of unmeasured ways; some students are extremely susceptible to friend influence while others are not. Moreover, the causal effect of, for example, friends’ average GPA likely varies across the vector of values; the effect of increasing friends’ average GPA from 2.0 to 3.0 is likely different than the effect of increasing friends’ average GPA from 3.0 to 4.0. The estimates presented in all of the instrumental variable analyses that follow are the weighted average of causal effects across both the heterogeneous population and the spectrum of values of the friend characteristics.

Of the four friend characteristics studied, only friends’ average self-reported GPA has a statistically significant impact on future GPA (p < 0.05). A one point increase in friends’ average GPA results in a 0.16 point increase in GPA at wave 2—less than the estimated friend effect from the OLS model. Given that a student takes six courses, this effect is approximately equivalent to increasing the grade in one course by one letter grade, from a C to a B for instance. This result falls in line with much of the research studying the relationship between friends and achievement, albeit with somewhat smaller effects. Accounting for the ways that students select into friendships and the fact that friends are simultaneously affecting one another, the average achievement level of friends plays a significant though modest role in affecting future grades.

This effect is comparable to the effects of many traditionally studied individual characteristics. For instance, an increase of one standard deviation in friends’ average GPA (0.52) increases GPA by 0.08 grade points. This effect is the same as the benefit an adolescent experiences by believing that college graduation is a certainty or by having a parent with some college education versus a parent who is a high school dropout.

By contrast, the average parental education of friends and their race/ethnic composition have no significant effect on future achievement. Whereas past research finds that parental education and, to a lesser extent, the race/ethnic composition of friends are associated with individual academic achievement (Choi et al. 2008; Fletcher...
et al. 2013; Riegle-Crumb and Callahan 2009), I find that these effects are nil, on average. Even when the academic composition of friends is not controlled, the race/ethnic composition and the parental education of friends have no significant effect on future GPA (see online supplement Table A.6).

Given that adolescents from different racial and ethnic backgrounds have friends with different academic characteristics, these patterns have the potential to contribute to achievement gaps in important ways. These effects, however, represent the average effect of friends' GPA across the analysis sample, assuming that friends equally influence educational behaviors. As I argue earlier, one of the many ways in which friend effects may vary across a population is by race/ethnicity. The next section investigates this possible heterogeneity in friend effects.

### Heterogeneous Friend Effects across Race/Ethnicity

Table 3 shows the results of the instrumental variable analysis predicting the effect of friends’ average GPA on achievement trajectories, separated by race/ethnicity. To ensure sufficient statistical power I combine the black and Latino students, estimating friend effects among white students and among black/Latino students. These models include all controls discussed earlier as well as school-by-grade fixed effects. Although race/ethnic heterogeneity in all friend effects was explored, in each case, the effect of parental education, proportion black, and proportion Latino were never statistically different from zero and did not differ across race and ethnicity.

I find marginally statistically significant friend effects for both white and black/Latino students. In both cases adolescents whose friends have higher GPAs are more likely to increase their achievement over time. However, these effects do not differ across race/ethnicity. Friend effects are equally strong among minority students as among white students. Overall, there is little evidence to suggest that the effect of friends’ average GPA on future GPA differs across race/ethnicity.

Although past research argues that friend effects should differ in important ways across race/ethnicity, similar to Crosnoe et al. (2003) I find no difference on average. Moreover, after accounting for selection and endogeneity, I am unable to confirm the results of Riegle-Crumb and Callahan (2009) that Latino adolescents, in particular, benefit from friendships with other Latinos or from friendships with students with highly educated parents.

### Linking Patterns of Friendship to Inequality

Given these results, I can assess the impact of different friendship choices on achievement differentials. What proportion of the racial/ethnic gap in academic performance can be explained by differences in the characteristics of adolescents’ chosen friends?

Table 4 shows the results of a simple decomposition of racial/ethnic differences in GPA. Given the estimated effect of friends’ academic achievement shown in Table 2 and the conclusions drawn from Table 3 that friend effects do not vary across race/ethnicity on average, the white-black and white-Latino differences in GPA can be decomposed to show how much of these differences are explained by differences in the characteristics of adolescents’ chosen friends. For example, if black students had friends with the GPA characteristics of white students (row 7: 2.89 instead of row 8: 2.67), the resulting difference in predicted GPA the following year for black students would be 0.04 grade points (row 5: (estimate of friend effect × white friends’ GPA)-(estimate of friend effect × black friends’ GPA)). This represents 19 percent of the GPA gap between black and white adolescents in the analysis sample (row 5/row 1). In other words, if black adolescents had friends with the achievement characteristics of white adolescents in this sample, the achievement gap between these groups would be 19 percent smaller. Similarly if Latino adolescents had friends with the same academic characteristics as white adolescents, the achievement gap would decrease by 17 percent.

Of course there are many things at play affecting friendship, including structural features of the network, patterns of homophily, and opportunity (Goodreau et al. 2009), as well as important dependencies between students. Simply changing the achievement level of friends is a somewhat simple-minded exercise, but it fulfills the goal of demonstrating the potential impact that friends...
Table 3: Parameter Estimates from IV Analyses Predicting Academic Achievement by Race/Ethnicity, Add Health

<table>
<thead>
<tr>
<th>Friend Characteristics</th>
<th>White</th>
<th>Black/Latino</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average GPA W1</td>
<td>0.179*</td>
<td>0.170*</td>
</tr>
<tr>
<td></td>
<td>(0.100)</td>
<td>(0.100)</td>
</tr>
<tr>
<td>Controls</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>School-Grade Fixed Effects</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>N</td>
<td>2,651</td>
<td>1,542</td>
</tr>
<tr>
<td>Statistical Difference Across Groups?</td>
<td>No</td>
<td></td>
</tr>
</tbody>
</table>

| * p < .1. |

Note: Standard errors in parentheses below unstandardized coefficients. Models account for clustering at the school level. Full results are available in Table A.7 in the online supplement.

Discussion

This article has two goals: 1) to provide a causal estimate of the effect that friends have on academic achievement and 2) to connect these causal effects of friends to patterns of racial and ethnic inequality in academic achievement. Using instrumental variable analysis and the characteristics of individuals’ friends’ friends as instruments to predict friends’ characteristics, I estimate friends’ impact on the academic achievement of youth. In contrast to past research, which finds that the parental education of adolescents’ friends and their racial/ethnic composition impacts achievement, I show that it is the achievement characteristics of youths’ friends that really matter. Although my estimate of the impact of friends’ academic achievement is somewhat smaller than that of past research, I show that adolescents with higher-achieving friends are more likely to increase their achievement compared to adolescents with lower-achieving friends. And in spite of strong arguments that these effects should differ across race and ethnicity, I find that the average effect of friends’ GPA on future GPA does not differ across race and ethnicity. Given these average effects, if black and Latino adolescents had friends with the achievement characteristics of white students, the achievement gap would be 17 to 19 percent smaller than it currently is. Although that is only a modest decline in the achievement gap, it represents an important and often overlooked source of difference among black and Latino youth.

As with all survey research claiming to estimate causal relationships, these results should be used cautiously and with acknowledgement of the assumptions necessary to assert causality. Ideally, I would use an instrument that is fully outside of the control and influence of youth—an act of mother nature that magically randomizes friendship ties. Unfortunately no such instrument exists. Instead, I make do with an imperfect but compelling instrument—the characteristics of adolescents’ indirect friends—and the untestable assumption that the characteristics of indirect friends are conditionally independent of the error term. If the characteristics of indirect friends are 1) directly affecting the future GPAs of respondents and/or 2) directly affected by the respondent, then this set of instruments is invalid.
I believe that the set of controls and school-by-grade fixed effects account for the major ways in which the set of instruments are correlated with the error term, but again, the assumption of independence cannot be tested.

Moreover, as I note earlier, this instrumental variable analysis captures the local average treatment effect of friends’ GPA, parental education, and racial/ethnic composition on a respondent’s future GPA. Individuals are likely heterogenous in the extent to which each is susceptible to friend influence and in the effect of friend characteristics across the spectrum of values. These estimates therefore represent the average of each individual’s causal response. The finding that friend effects do not differ across race/ethnicity, for example, does not mean that in some contexts or among some individuals these differences are not real, but rather that, on average, in this sample of students there are no racial/ethnic differences.

These results are also based on a selected sample in Add Health. The respondents in this analysis are higher-achieving, come from more advantaged backgrounds, and are more likely to have in-school friends compared to the larger representative Add Health sample. As a consequence, these results represent a conservative estimate of friends’ impact on inequality. The majority of excluded respondents are excluded because they do not nominate any in-school friends. If they truly have no friends, they miss out on the benefits of friendship, and the distance between the most and least advantaged widens. If these individuals have friends whom I do not observe but patterns of homophily hold and they are similarly affected by their friendships—as supplemental analyses suggest—then it again follows that friendship choices will continue to benefit the most advantaged and increase inequality.

In spite of these limitations, this analysis pushes us to think more carefully about the ways that friendships affect inequality. Inequality across race and ethnicity is not driven by the friend effect but rather by the differences in the characteristics of the friends made by black and Latino youth. Understanding how and why black and Latino youth make friends with different achievement characteristics, compared to white and Asian youth, is an important next step. Would changes in levels of segregation both in sociological science | www.sociologicalscience.com 272 July 2014 | Volume 1

Table 4: Percent of GPA difference explained by differences in friend GPA characteristics

<table>
<thead>
<tr>
<th></th>
<th>White-Black</th>
<th>White-Latino</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Difference in Self-Reported GPA</td>
<td>0.19</td>
<td>0.29</td>
</tr>
<tr>
<td><strong>Estimate of friend effect times</strong> . . .</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 White friends’ GPA</td>
<td>0.456</td>
<td>0.456</td>
</tr>
<tr>
<td>3 Black friends’ GPA</td>
<td>0.421</td>
<td></td>
</tr>
<tr>
<td>4 Latino friends’ GPA</td>
<td>0.408</td>
<td></td>
</tr>
<tr>
<td>5 Difference in effects</td>
<td>0.036</td>
<td>0.048</td>
</tr>
<tr>
<td>% of GPA difference explained</td>
<td>19%</td>
<td>17%</td>
</tr>
<tr>
<td>Lower bound</td>
<td>9%</td>
<td>8%</td>
</tr>
<tr>
<td>Upper bound</td>
<td>29%</td>
<td>25%</td>
</tr>
<tr>
<td>6 Estimate of friend effect</td>
<td>0.158</td>
<td></td>
</tr>
<tr>
<td>Friends’ Average GPA:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 White</td>
<td>2.89</td>
<td></td>
</tr>
<tr>
<td>8 Black</td>
<td>2.67</td>
<td></td>
</tr>
<tr>
<td>9 Latino</td>
<td>2.59</td>
<td></td>
</tr>
</tbody>
</table>

Note: Row 1 is calculated from the descriptives in Table 1. Row 2 is the product of rows 6 and 7. Row 3 is the product of rows 6 and 8. Row 4 is the product of rows 6 and 9. Row 5 is the difference between the values in rows 2 and 3, or 2 and 4. Row 6 is taken directly from Table 2, Model 3. Rows 7, 8, and 9 come from Table 1.
terms of race/ethnicity and socioeconomic status affect patterns of friendship, particularly the academic characteristics of black and Latino youths’ friends? Would changes in the size of a school or course-taking patterns create more opportunities for friendship across achievement levels? School composition and school organization—two characteristics that can be manipulated by policy—create opportunities for friendships and, potentially, opportunities to bridge gaps across race and ethnicity.

Inequality develops and is maintained through a complex set of processes. Although social science tends to place great emphasis on the school as both a source of inequality and an equalizer, the social relationships of children and adolescents within schools are often ignored. This analysis shows, controlling for all school characteristics that are constant across schoolmates, that friendships matter—but the characteristics that matter are limited. Ultimately it is the academic performance and the behaviors and norms established by the achievement behaviors of youths’ friends that contribute to gains in achievement. Whether these friends are black or white, have high-socioeconomic status or low socioeconomic status, does not matter, on average. This is in many ways a rosy story. Patterns of homophily across race and ethnicity, and to a lesser extent across socioeconomic status, are sticky. They are connected to dynamics of neighborhoods and families and likely more difficult to change than patterns of achievement homophily. Patterns of achievement homophily, by contrast, are likely more easily manipulated by changes in the organization of schools. Precisely how school organization and achievement homophily are related is the next question.

References


Alvarado, Steven E. and Ruth Lopez Turley. 2012. “College-Bound Friends and College Application Choices: Heterogeneous Effects for Latino and White Students.” *Social Science Research* 41:1451-1468. [http://dx.doi.org/10.1016/j.ssresearch.2012.05.017](http://dx.doi.org/10.1016/j.ssresearch.2012.05.017)


Calvo-Armengol, Antoni, Eleonora Patacchini, and Yves Zenou. 2009. “Peer Effects and Social Networks in Education.” *Review of Eco-


Acknowledgements: I would like to thank Robert Mare, Meredith Phillips, Renee Reichl Luthra, Rob Warren, and the UCLA Demography, Family, and Social Stratification Research Group for their advice and comments on previous drafts. Previous versions of this article were presented at the RC28 meeting at the University of Essex in 2011, the Population Association of America meetings in 2013, and the American Sociological Association meetings in 2013.

This research was supported by a National Institute of Child Health and Human Development traineeship through the California Center for Population Research and by a grant from the American Educational Research Association which receives funds for its “AERA Grants Program” from the National Science Foundation and the National Center for Education Statistics of the Institute of Education Sciences (U.S. Department of Education) under NSF Grant #DRL-0634035. Opinions reflect those of the author and do not necessarily reflect those of the granting agencies.

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 Eunice Kennedy Shriver 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). No direct support was received from grant P01-HD31921 for this analysis.

Jennifer Flashman: Center for Research on Educational Opportunity, University of Notre Dame. E-mail: Jennifer.A.Flashman.1@nd.edu.