Peer Influences on College Choices: New Evidence from Texas

Jason M. Fletcher[♦] [•]
Assistant Professor
Division of Health Policy and Administration
School of Public Health
Yale University
60 College Street, #303
New Haven, CT 06520
jason.fletcher@yale.edu
Fax: (203) 785 6287

Abstract:
In this paper, I use newly collected survey data from high school students in Texas to examine the importance of peer influences on college choices. In this survey, respondents (and their classmates) recorded their preferences for attending specific colleges, and a follow up survey recorded their college enrollment decisions. This paper uses this information to present the first empirical examination of whether individuals who report preferences for "unpopular" colleges are less likely to attend their preferred college. The rich data set allows the use of often unavailable information such as distance to college, and the construction of the "unpopularity" variable allows the use of school-level fixed effects. I also examine whether the influences of peer preferences vary by race and gender. Results indicate that individuals with 10 percentage points more classmates with matching college preferences are 3 percentage points more likely to enroll in their preferred college.

JEL Classification: I2, J24

Keywords: Demand for Schooling, Human Capital, Peer Effects, College Enrollment

---

[♦] I thank the Texas Higher Education Opportunity Project for generous financial support during this project. I also thank Marta Tienda, Steven Lehrer, Spencer Banzhaf, Gigi Foster, Andrew Grodner, Ethan Cohen-Cole, Sean Nicholson, and Suzanne Dancer for very helpful suggestions as well as participants of the Southern Economics Association Annual Meeting, the THEOP Summer Data Workshop, and the THEOP Research Conference at Princeton for helpful conversations on a previous version of this paper. Dawn Koffman provided very helpful assistance with the data. All errors are my own.
[•] This research uses data from the Texas Higher Education Opportunity Project (THEOP) and acknowledges the following agencies that made THEOP data available through grants and support: Ford Foundation, The Andrew W. Mellon Foundation, The William and Flora Hewlett Foundation, The Spencer Foundation, National Science Foundation (NSF Grant # SES-0350990), The National Institute of Child Health & Human Development (NICHD Grant # R24 H0047879) and The Office of Population Research at Princeton University

Introduction

There is a large literature that documents the importance of peer influences on adolescent choices. Many researchers cite the Coleman Report (1966) as the empirical starting point for evidence that peer effects in school settings contribute to achievement outcomes. More recently, there has been an increasing amount of research that attempts to document peer effects for educational and other outcomes. For example, peer decisions and/or peer characteristics have been shown to be important in predicting elementary school achievement (Hanushek et al. 2003, Hoxby 2000, Ammermueller and Pischke 2007, Lavy and Schlosser 2007, Cooley 2007), middle school achievement (McEwan 2003, Summers and Wolfe 1977, Lavy and Schlosser 2007), high school achievement (Ding and Lehrer 2007), and achievement during college (Sacerdote 2001, Zimmerman 2003, Fletcher and Tienda 2008).[1]

Besides academic achievement, peer influences have been shown to be important on several other educational outcomes, including enrollment in college (Fletcher 2008), choice of college major (Lyle 2007)[2], and whether to join a fraternity (Sacerdote 2001). Research examining the importance of peer influences has lagged behind other major educational decisions, including the choice of which college to attend. This omission is unfortunate due to the increasing importance of these choices for life outcome. For

---

[1] Peers have also been implicated in health outcomes, such as smoking, drinking, drug use, and sexual initiation (Gaviria and Raphael 2001, Kawaguchi 2004, Kremer and Levy 2003, Lundborg 2006, and Fletcher 2007 among others)

[2] Sacerdote (2001) finds no evidence of peer effects on choice of major. Arcidiacono and Nicholson (2005) find that evidence of peer influences on the choice of medical specialty disappear after controlling for school fixed effects.

example, Hoxby (2004) forcibly argues that the most important college decision has progressed from *whether* to attend college to *which* college to attend among the many alternatives. Data limitations and empirical difficulties in estimating the importance of peer influences on individual choices are likely two principal reasons for the limited research in this area.

In this paper, I use newly collected survey data from high school students in Texas to examine the importance of peer influences on college choices. In this survey, respondents (and their classmates) recorded their preferences for attending specific colleges, and a follow up survey recorded their college enrollment decisions. Therefore, the data indicate whether each individual enrolled in his/her preferred college as well as classmates' preferences for individual colleges. The extent that an individual's preferred college differs from the preferences of his/her peers is an indication of whether the individual has "unpopular" college preferences.

The empirical question examined in this paper is whether individuals who had stated preferences to attend "unpopular" colleges were less likely to attend their preferred college. There are multiple hypotheses that would imply that peer preferences may influence individual choices. One hypothesis is that peer preferences increase an individual's information about colleges. For example, an individual who plans to apply to a "popular" college may attain information from classmates about the college (visit day information, reminders about important deadlines, etc.) than individuals who plan to apply to "unpopular" colleges. Alternatively, adolescents may receive utility from conforming with "popular preferences" (e.g. Bernheim 1994), for example, from an increase in discussions with classmates about their shared interest in a specific college.

The focus of this paper is whether peer preferences are associated with an individual's college choice rather than establishing the mechanism.

Although assessing causal mechanisms is beyond the scope of this paper, the dataset does allow the use of often unavailable information such as distance to college and contains a relatively rich set of covariates. Additionally, the construction of the "unpopularity" variable used in this paper allows the use of school fixed effects. Results indicate that individuals with 10 percentage points more classmates with matching college preferences are 3 percentage points more likely to enroll in their preferred college. The association is most important for Hispanic students in Texas and non-existent for black students.

Background

There are large literatures on the determinants of both *whether* to enroll in college and *which* college to enroll in. Not surprisingly, it has been shown that more academically able students and students from families with greater resources are more likely to pursue post-secondary education (McDonough 1997, Cameron and Heckman 1998, 2001). There is also some evidence that going to high school with more college-going classmates increases an individual's propensity to attend college (Fletcher 2008).

As mentioned above, there has recently been a push within economics to focus more attention on examining why individuals choose specific colleges rather than their decision to enroll in any college (Hoxby 2004). Much of the literature on college choice has focused on pecuniary returns to attending a selective college in comparison to a less selective college (Brewer et al. 1999, Loury and Garman 1995). However, there is also evidence that the choice of which college to attend is driven by multiple factors in addition to college selectivity and quality. The desirability of having a good "match" between a student's ability and college peer ability has been shown to be important (Fuller, Manski and Wise 1983, Light and Strayer 2000). Students also are sensitive to the costs of attending colleges, both in terms of geographic distance (Turley 2006, Fletcher 2008), and in terms of financial aid offerings (Fuller, Manski and Wise 1983, Avery and Hoxby 2004).

There is also reason to believe that non-pecuniary factors are an important aspect of college choice for many students. For example, family traditions (i.e. legacies) of attending specific colleges likely motivate some individuals in their college choices. Success of college athletic teams has also been shown to increase college applications (Pope and Pope 2007).

Also potentially important are the preferences of individuals in one's peer group. The principal goal of this paper is to examine the importance of this peer influence. In particular, are adolescents with "unpopular" preferences for specific colleges less likely to attend their preferred college? While there have been no attempts to estimate the importance of peer preferences on college choice in the literature, there are several papers that examine similar questions, including choice of medical specialty and choice of college major.

A first paper that examines peer influences on choices during college is Lyle (2007). The author examines whether "role model effects" are important in predicting choice of college major at West Point. In particular, the author estimates whether a freshman's eventual choice of major is associated with the proportion of upperclassmen in his/her randomly assigned peer group (i.e. company). He finds suggestive evidence of a relationship for freshman who major in engineering, social sciences, and natural sciences (although only engineering is statistically significant) but negative associations in other majors. In contrast, Sacerdote (2001) finds no evidence that randomly assigned roommates at Dartmouth College influence the choice of college major.

Arcidiacono and Nicholson (2005) used the universe of medical students at US medical schools across three cohorts to examine whether individuals switch their preferred area of specialty because of peer group influences. The authors' measure of peer preferences was the proportion of an individual's medical school cohort who indicated a preference for a high-income medical specialty in their first year of medical school.[3] They find evidence that this measure of peer preferences is associated with an individual's medical specialty choice in their baseline specifications, but this association disappears after controlling for school-level fixed effects. One drawback of their measure is that it likely correlates with peer quality, and since available positions for high income medical specialties are rationed, it is not clear how peer preferences for medical specialties are influencing their classmates. Having a greater number of peers who prefer high income specialties could (1) increase a classmate's propensity to choose a high income specialty through learning about the advantages of choosing these specialties or (2) decrease a classmate's propensity to choose a high income specialty because of the finite number of available slots.

[3]High-income specialties included surgery, medical sub-specialties, radiology, anesthesiology, pathology, and obstetrics. Low-income specialties included internal medicine, emergency medicine, pediatrics, family practice, and psychiatry.

In contrast to the Arcidiacono/Nicholson measure, I create a measure in the college choice context that indicates the proportion of peers who prefer the same college.[4]

This measure abstracts from the quality of the preferred college to examine the average influence of peer preferences on college choice, although I control for the selectivity of a student's preferred college.

As an example, assume there are four students in a Texas high school who report college preferences. Student A prefers attending the University of Texas-Austin, Student B prefers attending Community College X, and Students C and D prefer attending Texas A & M. In this case, Students A and B both have identical proportions of classmates with the same college preference (0%) even though they prefer colleges of differing quality, while Students C and D also have the same proportion (33%). In addition to abstracting from school quality, the measure of "unpopularity" of college preferences used in this paper also allows school fixed effects to be controlled, because there is within-school variation in the proportion of classmates with the same college preference (as the above example shows). Finally, because the analysis does not regress classmate *outcomes* on individual *outcomes*, the well known issue of the reflection problem (Manksi 1993) in empirical models of peer effects is not directly applicable. In

[4] A similar measure to Arcidiacono/Nicholson in the college choice context would be to measure the proportion of high school classmates who prefer to attend an elite university. The potential disadvantages with this measure are similar to the medical specialty measure described above—it likely measures peer quality, learning opportunities about elite colleges, as well as fewer available slots at specific elite colleges. Since the dataset I use is of a single cohort of students, I am unable to use the econometric approach of Arcidiacono and Nicholson.

particular, a student's choice to attend a specific college after high school graduation can not affect his classmates' stated preferences for particular colleges elicited during high school—there is no direct reflection problem. An empirical issue, though, that reduces the ability to pin down causal estimates is that preferences for colleges that are formed during high school are potentially jointly determined among classmates. Student A's strong preference to attend UT-Austin may influence Student B's preference formation and also influence the likelihood of Student A attending their preference college. Since these data do not have measures capturing the strength or dynamics of preference formation among classmates, the results must be viewed as suggestive of peer influence rather than conclusive. On the other hand, a strength of the approach is to limit a large set of factors that determine preference formation for classmates, such as distance to a specific college, guidance counselor assistance, and other shared high school factors, through the use of school fixed effect controls.

Data

This paper uses the newly available Texas Higher Education Opportunity Project (THEOP) data.[5] The THEOP data is a multi-year study that began in fall 2000. In addition to gathering administrative data from 10 colleges and universities in Texas, the centerpiece of the study is a two-cohort longitudinal survey of sophomores and seniors who were enrolled in Texas public schools in spring 2002. This paper will focus on the senior cohort.[6] The baseline survey (Wave 1) was conducted on a stratified random

[5] Complete information can be found at http://www.texastop10.princeton.edu/index.html

[6] The sophomore cohort was not followed post-high school.

sample of 105 public high schools in the state of Texas and consists of 13,803 seniors. The baseline survey asked students about their course taking and grades, college perceptions, future plans and demographic information, including race, family background, and household structure. Seniors were asked a battery of questions about college preferences, the colleges applied to, and plans to attend college. A random sample of 5,836 respondents from the senior cohort were re-interviewed (Wave 2) one year after graduating from high school to ascertain primary post-secondary school activity, military enlistment, labor force participation, etc. [7]

This paper uses the combined Wave 1 and Wave 2 data. As mentioned above, the high school seniors during Wave 1 were asked to provide the names of the colleges that they preferred to attend. While the focus of the paper is necessarily on the 5,836 individuals who were followed in Wave 2 and have information on college outcomes, this paper also uses the Wave 1 survey to construct various aspects of each individual's high school environment using information from individuals who were not followed in Wave 2 but nonetheless provide details on preferences for college and other school-level characteristics.

As noted above, the sample size starts at 5,836 individuals. Since several variables will be created at the school-level for each senior class within each school, I drop sixteen individuals who are sampled in schools with fewer than ten other students. Non-response for gender and race forces the deletion of six hundred individuals so that 5,224 individuals remain. Another one hundred and thirty individuals are dropped because of unreported grades during high school. I use single-imputation methods to estimate mother's education level for three hundred and fifty individuals, leaving a

[7] A public version of the data is available at http://opr.princeton.edu/archive

sample size of 5,095. Finally, I drop individuals who did not report a college preference, leaving approximately 3,600 students for the analysis sample. It is important to note that the use of school fixed effects in the empirical analysis will control for the proportion of individuals with unknown or no college preferences in each high school, since these measures do not vary within high school.

Summary statistics are presented in Table 1 below. Nearly 85% of the individuals in the analysis sample of 12th graders report some post-secondary experience[8] by Wave 2; 46% enrolled in the preferred school. Unfortunately, there is relatively little information in the survey that captures family resources (e.g. income) at the individual level, so mother's education attainment is used as an indicator of income. At the school level, I include senior enrollment, the proportion male, the proportion Hispanic, the proportion black, and the proportion who are economically disadvantaged. Additionally, the THEOP data contains measures of the distance from each high school to each college in the state, which allows inclusion of an important set of variables that are related to college decisions by adolescents but often neglected in previous research.[9] Finally, the proportion of each individual's classmates (in the Full Wave 1 Sample of 13,000 who reported college preferences) who share his/her preference for specific colleges is calculated. On average, nearly 11% of an individual's classmates prefer the same college.

[8] This includes vocational, technical, or trade school, and those who have taken courses from a university or college for academic credit. For brevity, I refer to all these institutions as "college" throughout the paper.

[9] Fletcher (2008) and Turley (2006) are recent exceptions.

Empirical Methods and Results

The empirical methodology employed in this paper proceeds in two steps. First, I examine the association between individual and school-level variables and the proportion of classmates with the same college preferences. This empirical exercise is performed to examine potential selection into high schools based on heterogeneity in college preferences (which is the variable of interest in step two). The second empirical step will be to estimate the correlation between whether an individual enrolls in his/her preferred college and the proportion of classmates who also preferred the college—the measure of having "popular" preferences for college vis-à-vis one's high school classmates.

The association between matching preferences and individual and school-grade-level characteristics is estimated using OLS regression of the following form:

$$\%Match_{is} = \beta X_{is} + \theta \overline{X}_{-is} + \varepsilon_{is} \tag{1}$$

where $\%Match$ indicates the proportion of classmates with the same preference for a particular college as student *i* in high school *s* (recall, all individuals are in the 12th grade). Characteristics such as race, gender, grade point average, and maternal education are measured at the individual and school-grade-level, where $\overline{X}_{-is}$ indicates the average characteristic (except the focal individual) at the school-grade level. The individual error term is allowed to be arbitrarily correlated for individuals in the same high school/grade.

If the proportion of classmates who share the same college preference is associated with unobservable high school quality, then we may worry that the "match" measure may be picking up the effects of high school quality on college choices in step two. In fact, the results in Table 2 show that individuals with more highly educated mothers attend school with *fewer* classmates with the same preferences for colleges. The results also indicate that black students attend high schools with 3% fewer classmates who share the same college preferences as white students and Hispanic students attend high schools with a 1% higher proportion than white students. Students in schools with poorer student-bodies have a higher proportion of classmates with identical college preferences, as do students who attend schools with lower proportions of minority students. Interestingly, an individual's grade point average is not associated with having a higher proportion of classmates with the same college preferences. These results suggest that while the proportion of an individual's classmates who share his/her preference for college likely represents heterogeneity in the student population (e.g. along racial/income lines), it does not seem to be correlated with measures of high school quality that individuals with higher ability or greater family resources would select into.

The results may seem counterintuitive, but the reader is reminded that the $\%Match$ variable is not necessarily related to quality of preferred college. Students who report a desire to attend a low-quality community college but whose preferences match those of their classmates can have a high value of $\%Match$. Likewise, students in schools where all seniors prefer to attend an elite university, but whose preferences differ in *which* elite university to attend (Harvard vs. Yale vs. Princeton, etc.) can have a low value of $\%Match$.

Next, I examine the determinants of whether individuals make the following post-high school choices: attend their preferred college, attend a non-preferred college, and attend no college. I estimate a multinomial logistic regression specification of

$$Outcome_{is} = \%Match_{is} + \beta X_{is} + \theta \overline{X}_{-is} + \lambda W_s + \delta C_{is} + \varepsilon_{is} \qquad (2)$$

where $Outcome$ is the post-secondary choice of student *i* in high school *s* and $\%Match$ is the proportion of students in the same high school-grade who report the same college preference as student *i*. As before, the X vector measures individual/family level characteristics and the $\overline{X}_{-is}$ is a measure of the school-grade average of the individual characteristic in X (except the focal individual). Additionally, in some specifications I include school-level characteristics such as distance to nearest college (W) and also characteristics of the individual's preferred college, such as selectivity and costs of attendance (C_{is}).[10] All coefficients in Table 3 represent marginal effects and should be interpreted relative to the omitted category of enrolling in the preferred college; robust standard errors are included in parentheses.

The results suggest that a student with a 10 percentage point greater number of classmates who share the same college preference is associated with a 6 percentage point increase in the probability of attending a student's preferred college—recall that 45% of the students in the sample attend their preferred college. The magnitude of this association is similar to an increase in GPA of nearly one point. The results also show that minority students are less likely to attend their preferred college than white students, with decreases of 5, 12, and 12 percentage points for blacks, Hispanics, and students of other races, respectively. Additionally, students who attend schools with higher minority shares are more likely to attend a non-preferred college.

While the results for %Match in columns 1 and 2 are interesting, one limitation is that some students may "over reach" when indicating their preferred college, so that one

[10] The data on selectivity and costs of attendance come from the Integrated Postsecondary Education Data System (IPEDS) and were generously provided by Marta Tienda.

reason they attend a non-preferred college is that they were too optimistic in indicating their preferred college. In this case, the coefficient on %Match is picking up students whose choices may have been changed by peer influences as well as students who indicated preferences for colleges they could not attend—because they were academically ineligible, faced significant borrowing constraints, or other reasons. In columns 3-6, I control for measures of college selectivity and costs of the preferred college as well as distance to the nearest college in order to control for "over reaching" students and heterogeneity in the attributes of preferred colleges. Adding these variables decreases the magnitude of the association between %Match and enrolling in a non-preferred college by 50%. The results suggest that a student with a 10 percentage point greater number of classmates who share the same college preference is associated with a 3 percentage point increase in the probability of attending a student's preferred college. This is still a non-trivial difference in the probability of attending one's preferred college—comparable with nearly a 1/3 point increase in GPA.

While previous specifications capture a variety of school and individual level factors, there may be a concern that the %Match variable may be correlated with other school-quality or other confounding influences not yet controlled in the specifications. Next, to reduce this concern, results in Table 4 take advantage of the construction of the %Match variable, which varies within high-schools by using school-level fixed effects. Results without fixed effects are repeated in columns 1 and 2. The results with fixed effects eliminate the association between %Match and the probability of attending no college, but the association between %Match and attending a non-preferred college remain unchanged. Broadly speaking, the inclusion of school-level fixed effects does not qualitatively change many of the individual-level variables.

Finally, as there is compelling evidence that there are racial differences in college preferences (e.g. Niu et al. 2006), I estimate equation (2) separately for white, black, and Hispanic students in Table 5. The results indicate that Hispanic students are most affected by the presence of peers with the same preference for colleges. The results suggest that a Hispanic student with a 10 percentage point greater number of classmates who share the same college preference is associated with a 7 percentage point increase in the probability of attending a student's preferred college. In contrast, there is no evidence that black students make college choices based on peer preferences. Additionally, the number of siblings decreases the likelihood of enrolling in one's preferred college only for Hispanic students. The cost of attending one's preferred college is positively associated with not attending college for minority students only.[11]

Conclusion

The main results in this paper suggest that peer preferences in high school are associated with an individual's eventual choice of which college to attend for a recent sample of high school seniors from Texas. In particular, an individual in a high school with 10 percentage point more peers with matching preferences for a particular college is 3 percentage points more likely to attend his/her preferred college. This association is

[11] Results were similar between males and females and are available upon request. Results were the %Match variable was calculated for own-race or own-gender were also very similar to the results presented in the paper and are also available upon request.

robust to including high school fixed effects as well as controlling for attributes of the individual's preferred college, such as selectivity and cost. These results suggest that an important aspect of college choice for adolescents is peer influence. Since the choice of which college to attend has become more central to research in the economics of higher education (Hoxby 2004), the likelihood that peers influence this choice is important to document and examine in greater detail. If there are important peer influences on an individual's college choice, policies that shape an individual's peer environment may also shape individual college choices. This implication could be particularly relevant for increasing the representation of minority students and students from areas with low college-going traditions into certain colleges (e.g. selective college, geographically diverse colleges, etc).

While this paper provides evidence of the importance of peer preferences on college choices, there are several alternative hypotheses that could be consistent with this finding; each likely implies different forms of intervening policies. It could be the case that peers' preferences for certain colleges provide important information in one's college decision. This situation may imply that students are actually making superior choices in not enrolling in the college they reported to prefer. The information provided by students about alternative colleges could lead to more informed and better choices. This hypothesis likely suggests that further infusions of information about college choices could be the primary relevant policy in taking advantage of peer preferences. Alternatively, peers' preferences could instill a social norm of "acceptable choices" within high school peer groups. In this case, peers may constrain the college options for an individual (see Niu and Tienda forthcoming), and individuals who choose non-preferred colleges for social reasons may be making sub-optimal choices. In this case, efforts to elevate the norm of which types of colleges are “appropriate” and socially valued choices within high schools may be needed.

References

Ammermueller, Andreas and Jorn-Steffen Pischke. (2007). "Peer Effects in European Primary Schools: Evidence from PIRLS." Working Paper

Arcidiacono, Peter. (2004). "Ability Sorting and Returns to College Major." *Journal of Econometrics*, 121: 343-375

Arcidiacono, Peter and Sean Nicholson. (2005). "Peer Effects in Medical School." *Journal of Public Economics*, Vol 89, pp. 327-350

Avery, Christopher and Caroline Hoxby. (2004). "Do and Should Financial Aid Packages Affect Students' College Choices?" In *College Choices: The Economics of Which College, When College, and How to Pay for It*. University of Chicago Press

Bernheim, B. Douglas. (1994). "A Theory of Conformity." *Journal of Political Economy*, Vol 102 (5): 841-877

Brewer, Dominic, Eric Eide, and Ronald Ehrenberg. (1999). "Does it Pay to Attend an Elite Private College? Cross Cohort Evidence on the Effects of College Quality on Earnings." *Journal of Human Resources*, Vol 34 (1): 104-123

Cameron, Stephen and James Heckman. (1998). “Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males.” *Journal of Political Economy*, Vol 106 (21)

Cameron, Stephen and James Heckman (2001). “Dynamics of Educational Attainment for Black, Hispanic, and White Males.” *Journal of Political Economy,* Vol 109 (31)

Coleman, James et al. (1966). *Equality of Educational Opportunity* . Washington, DC: United States Government Printing Office

Cooley, Jane. (2007). “Desegregation and the Achievement Gap: Do Diverse Peers Help?” University of Wisconsin-Madison Working Paper

Ding, Weili and Steven Lehrer (2007). “Do Peers Affect Student Achievement in China’s Secondary Schools?” *Review of Economics and Statistics*.89 (2)

Fletcher, Jason M. (2008). “College Enrollment and Social Interactions: Evidence from NELS.” Yale University Working Paper

Fletcher, Jason M. (2007). “Social Multipliers in the Sexual Initiation Decisions among U.S. High School Students.” *Demography*

Fletcher, Jason M. and Marta Tienda (2008). “High School Peer Networks and College Success: Lessons from Texas.” Yale University Working Paper

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

Fuller, Winship, Charles Manski, and David Wise. (1983). “New Evidence on the Economic Determinants of Postsecondary Schooling Choices.” *Journal of Human Resources*, 17 (4): 477-498

Gaviria, Alejandro and Steven Raphael. (2001) “School-Based Peer Effects and Juvenile Behavior.” *Review of Economics and Statistics*, Vol 83 (2)

Hanushek, E.A., Kain, J.F., Markman, J.M., Rivkin, S.G., (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics* 18 (5), 527–544.

Hoxby, Caroline. (2000). “Peer Effects in the Classroom: Learning from Gender and Race Variation.” NBER Working Paper 7867

Hoxby, Caroline. (2004). “Introduction.” in *College Choices: The Economics of Which College, When College, and How to Pay for It*. University of Chicago Press

Lavy, Victor and Analia Schlosser. (2007). "Mechanisms and Impacts of Gender Peer Effects at School." Working Paper

Light, Audrey and Wayne Strayer. (2000). Determinants of College Completion: School Quality or Student Ability?" *Journal of Human Resources*, Vol 32 (2)

Loury, Linda and David Garman. (1995). "College Selectivity and Earnings." *Journal of Labor Economics*, Vol 13 (2)

Lundborg, Peter. (2006). "Having the Wrong Friends? Peer Effects in Adolescent Substance Use." *Journal of Health Economics*, Vol 25: 214-233

Lyle, David, (2007). "Estimating and Interpreting Peer and Role Model Effects from Randomly Assigned Social Groups at West Point." *Review of Economics and Statistics*, 89(2): 289-299

Manski, Charles and David Wise. (1983). *College Choice in America*. Harvard University Press: Cambridge, Mass.

McDonough, Patricia. (1997). *Choosing Colleges: How Social Class and Schools Structure Opportunity*. State University of New York Press: Albany

Niu, Sunny and Marta Tienda (forthcoming). “Choosing Colleges: Identifying and Modeling Choice Sets.” *Social Science Research*

Niu, Sunny, Marta Tienda, and Kalena Cortes. (2006). “College Selectivity and the Texas Top 10% Law.” *Economics of Education Review*, 25: 259-272

Pope, Devin and Jaren Pope. (2007). “Understanding College Choice Decisions: How Sports Success Garners Attention and Provides Information.” Working Paper

Pope, Devin and Jaren Pope (2007). “The Impact of Sports Success on the Quantity and Quality of Student Applications.” Working Paper

Sacerdote, Bruce. (2001). “Peer Effects with Random Assignments: Results from Dartmouth Roommates.” *Quarterly Journal of Economics*, 116, 2

Turley, Ruth Lopez. (2006). "College Proximity: Mapping Access to Opportunity.” University of Wisconsin-Madison Working Paper

Zimmerman D. (2003). “Peer Effects in Academic Outcomes: Evidence from a Natural Experiment.” *Review of Economics and Statistics*. 85 (1)

Tables

Table 1
Summary Statistics for THEOP Data
Approximate N=3,600

Variable	Obs	Mean	Std. Dev	Min	Max
College	3587	0.85	0.36	0.0	1.0
Cost of Preferred College ($10,000s)	3505	1.13	0.70	0.0	3.9
Selectivity of Preferred College (0=Non-Competitive, 9=Most)	3557	3.27	2.44	0.0	9.0
Attend Preferred College	3587	0.46	0.50	0.0	1.0
% Match	3587	10.74	12.95	0.0	60.6
Male	3587	0.41	0.49	0.0	1.0
White	3587	0.42	0.49	0.0	1.0
Black	3587	0.18	0.38	0.0	1.0
Hispanic	3587	0.29	0.45	0.0	1.0
Other Race	3587	0.11	0.31	0.0	1.0
Grade Point Average	3587	3.23	0.64	1.0	4.0
Years of Maternal Education^	3587	13.97	2.66	0.0	19.0
Number of Siblings	3587	2.72	2.30	0.0	12.0
School-Grade Level Characteristics					
% Male	3587	47.23	4.69	30.0	66.7
Mean Years of Maternal Education	3587	15.70	2.22	12.5	26.6
% Black	3587	13.58	15.99	0.0	96.5
% Hispanic	3587	27.55	26.74	0.0	89.3
% Other Race	3587	9.41	9.03	0.0	32.2
Log Enrollment	3587	5.91	0.89	0.0	7.4
% Disadvantaged	3587	31.95	22.89	0.9	93.8
Nearest 4 Yr College (Miles)	3587	11.20	11.08	0.7	83.3
Nearest 2 Yr College	3587	13.05	15.00	0.5	74.3
Nearest Private College	3587	47.85	64.73	0.7	247.9
Nearest College	3587	6.73	8.24	0.5	60.2

^Imputed Variable, Sample includes individuals who reported a preferred college in Wave I and were sampled in Wave II

Notes: %Match variable: 21% of the sample has no classmates with matching college preferences, the median is 6%. There were 400 different colleges nominated as a preferred college.

Table 2
Individual and School-Level Determinants of Proportion of Peers with the Same Preference for College

Outcome	% Match	% Match
Method	OLS	
Fixed Effects	No	School
Male	-0.709	-0.495
	(0.521)	(0.494)
Black	-2.809***	-1.879***
	(0.580)	(0.637)
Hispanic	1.236**	1.363*
	(0.604)	(0.759)
Other Race	1.901*	2.235**
	(1.009)	(1.007)
Maternal Education	-0.181*	-0.170**
	(0.095)	(0.078)
GPA	0.252	0.090
	(0.386)	(0.412)
Number Sampled	0.007	
	(0.010)	
Mean Maternal Education	0.140	
	(0.240)	
% Male	0.068	
	(0.125)	
% Black	-0.137***	
	(0.046)	
% Hispanic	-0.088	
	(0.060)	
% Other Race	0.081	
	(0.080)	
Log (Enrollment)	-0.965	
	(0.863)	
% Disadvantaged	0.190**	
	(0.077)	
Constant	9.740	12.724***
	(9.209)	(1.788)
Observations	3587	3587
R-squared	0.075	0.012
Number of Schools		88

Robust standard errors in parentheses, *** 1%, **5%, *10%
Errors are clustered at the school level

Table 3
Determinants of College Choices:
Multinomial Logistic Regression Results

Outcome	No College	Non Preferred College	No College	Non Preferred College	No College	Non Preferred College
Method	MNL	MNL	MNL	MNL	MNL	MNL
Fixed Effects	No	No	No	No	No	No
% Match	0.001**	-0.006***	0.002***	-0.004***	0.001***	-0.003***
	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)
GPA	-0.084***	-0.071***	-0.076***	-0.085***	-0.075***	-0.083***
	(0.008)	(0.014)	(0.008)	(0.015)	(0.008)	(0.015)
Black	-0.009	0.051*	0.001	0.062**	-0.001	0.057**
	(0.019)	(0.028)	(0.019)	(0.029)	(0.019)	(0.029)
Hispanic	0.057***	0.117***	0.061***	0.110***	0.061***	0.109***
	(0.017)	(0.028)	(0.017)	(0.028)	(0.017)	(0.029)
Other Race	-0.033	0.123***	-0.026	0.114***	-0.025	0.109***
	(0.024)	(0.030)	(0.024)	(0.031)	(0.024)	(0.031)
Number of Siblings	0.012***	0.001	0.011***	0.003	0.011***	0.002
	(0.002)	(0.004)	(0.002)	(0.004)	(0.002)	(0.004)
Maternal Education	-0.007***	-0.002	-0.006***	-0.004	-0.006***	-0.004
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
Mean Maternal Education	-0.005**	-0.003	-0.004*	-0.006	-0.004	-0.003
	(0.002)	(0.004)	(0.002)	(0.004)	(0.002)	(0.004)
% Black	-0.001**	0.002**	-0.001**	0.002**	-0.001**	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
% Hispanic	-0.001	0.002	-0.001	0.002	-0.001	0.003
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)
% Other Race	0.001	0.004***	0.001	0.004***	0.001	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
% Disadvantaged	0.003***	-0.002**	0.002***	-0.002**	0.003***	-0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
College Selectivity			-0.012***	0.009**	-0.012***	0.008*
			(0.003)	(0.004)	(0.003)	(0.004)
Cost of Attendance			0.038***	0.051***	0.036***	0.058***
			(0.010)	(0.015)	(0.010)	(0.015)
Distance to 4 Yr					-0.002***	0.003***
					(0.001)	(0.001)
Distance to 2 Yr					0.001***	-0.002***
					(0.000)	(0.001)
Distance to Private					-0.000	-0.001**
					(0.000)	(0.000)
Observations	3587	3587	3505	3505	3505	3505

Outcomes include (1) No College, (2) Non-Preferred College, and (3) Preferred College (omitted). Marginal Effects evaluated at the mean reported. Robust standard errors in parentheses, *** 1%, **5%, *10%. Errors are clustered at the school level
Distance variables: distance in miles from the individual's high school to the nearest (1) four year college (2) two year college, and (3) private college
Additional control variables: male, school-level % male, log (enrollment), Constant

Table 4
Determinants of College Choices:
Multinomial Logistic Regression Results
Including School-Level Fixed Effects

Outcome	No College	Non Preferred College	No College	Non Preferred College
Method	MNL	MNL	MNL	MNL
Fixed Effects	No	No	School	School
% Match	0.001***	-0.003***	0.000	-0.004***
	(0.000)	(0.001)	(0.000)	(0.001)
GPA	-0.075***	-0.083***	-0.050	-0.099***
	(0.008)	(0.015)	(0.033)	(0.025)
Black	-0.001	0.057**	0.006	0.064**
	(0.019)	(0.029)	(0.012)	(0.030)
Hispanic	0.061***	0.109***	0.048	0.114***
	(0.017)	(0.029)	(0.033)	(0.036)
Other Race	-0.025	0.109***	-0.007	0.104***
	(0.024)	(0.031)	(0.014)	(0.035)
Number of Siblings	0.011***	0.002	0.006	0.004
	(0.002)	(0.004)	(0.004)	(0.004)
Maternal Education	-0.006***	-0.004	-0.004	-0.005
	(0.002)	(0.003)	(0.003)	(0.003)
Mean Maternal Education	-0.004	-0.003		
	(0.002)	(0.004)		
% Black	-0.001**	0.003***		
	(0.001)	(0.001)		
% Hispanic	-0.001	0.003		
	(0.001)	(0.002)		
% Other Race	0.001	0.005***		
	(0.001)	(0.001)		
% Disadvantaged	0.003***	-0.002**		
	(0.001)	(0.001)		
College Selectivity	-0.012***	0.008*	-0.006	0.005
	(0.003)	(0.004)	(0.004)	(0.004)
Cost of Attendance	0.036***	0.058***	0.021	0.062***
	(0.010)	(0.015)	(0.015)	(0.019)
Distance to 4 Yr	-0.002***	0.003***		
	(0.001)	(0.001)		
Distance to 2 Yr	0.001***	-0.002***		
	(0.000)	(0.001)		
Distance to Private	-0.000	-0.001**		
	(0.000)	(0.000)		
Observations	3505	3505	3505	3505

Outcomes include (1) No College, (2) Non-Preferred College, and (3) Preferred College (omitted). Marginal Effects evaluated at the mean reported. Robust standard errors in parentheses, *** 1%, **5%, *10%. Errors are clustered at the school level.
Distance variables: distance in miles from the individual's high school to the nearest (1) four year college (2) two year college, and (3) private college
Additional control variables: male, school-level % male, log (enrollment), constant

Table 5
Determinants of College Choices:
MNL Results Separated by Race

Outcome	No College	Non Preferred College	No College	Non Preferred College	No College	Non Preferred College
Method	MNL	MNL	MNL	MNL	MNL	MNL
Fixed Effects	No	No	No	No	No	No
Sample	White	White	Black	Black	Hispanic	Hispanic
% Match	0.001***	-0.005***	0.001	0.001	0.002**	-0.007***
	(0.000)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
GPA	-0.049***	-0.116***	-0.067***	-0.048	-0.112***	-0.024
	(0.009)	(0.021)	(0.024)	(0.037)	(0.024)	(0.028)
Male	0.021*	0.000	-0.014	0.052	0.030	-0.012
	(0.012)	(0.026)	(0.030)	(0.045)	(0.029)	(0.034)
Number of Siblings	0.007***	0.003	0.017***	-0.011	0.017***	0.012*
	(0.003)	(0.006)	(0.005)	(0.009)	(0.005)	(0.007)
Maternal Education	-0.010***	-0.018***	-0.026***	0.012	-0.002	-0.001
	(0.003)	(0.006)	(0.007)	(0.010)	(0.004)	(0.005)
Mean Maternal Education	0.000	0.001	-0.001	-0.013	-0.006	-0.007
	(0.004)	(0.008)	(0.006)	(0.009)	(0.005)	(0.007)
% Male	0.001	0.002	-0.002**	0.005***	0.001	0.004**
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
% Black	0.000	0.002	-0.003*	0.003	-0.001	0.002
	(0.001)	(0.002)	(0.002)	(0.003)	(0.001)	(0.001)
% Hispanic	0.001	0.002	-0.001	-0.000	-0.004	0.004
	(0.001)	(0.003)	(0.004)	(0.006)	(0.003)	(0.004)
% Other Race	0.000	0.001	-0.001	0.006**	0.005*	0.006**
	(0.001)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)
Log (Enrollment)	-0.026***	-0.054**	0.003	-0.005	-0.005	-0.005
	(0.010)	(0.024)	(0.015)	(0.025)	(0.021)	(0.026)
% Disadvantaged	0.000	-0.003	0.004**	-0.002	0.004***	-0.002
	(0.001)	(0.002)	(0.002)	(0.003)	(0.001)	(0.002)
College Selectivity	-0.011***	0.013**	-0.012	0.012	-0.008	0.020**
	(0.003)	(0.006)	(0.009)	(0.013)	(0.008)	(0.008)
Cost of Attendance	0.013	0.024	0.044	0.138***	0.064**	0.003
	(0.012)	(0.021)	(0.030)	(0.046)	(0.025)	(0.029)
Nearest College	-0.002***	-0.003	0.000	0.001	0.001	0.001
	(0.001)	(0.002)	(0.003)	(0.005)	(0.002)	(0.002)
Constant	0.250**	0.761***	0.395	-0.036	0.262	0.042
	(0.115)	(0.265)	(0.242)	(0.372)	(0.212)	(0.260)
Observations	1472	1472	640	640	998	998

Outcomes include (1) No College, (2) Non-Preferred College, and (3) Preferred College (omitted).
Marginal Effects evaluated at the mean reported. Robust standard errors in parentheses, *** 1%, **5%, *10%. Errors are clustered at the school level
Distance variable: Minimum of the distance in miles from the individual’s high school to the nearest (1) four year college (2) two year college, and (3) private college