WISCAPE
WORKING PAPER

Interpreting Community College Effects in the Presence of Heterogeneity and Complex Counterfactuals

Jennie E. Brand
University of California–Los Angeles

Fabian T. Pfeffer
University of Michigan

Sara Goldrick-Rab
University of Wisconsin–Madison

April 2012

Address correspondence to:

Jennie E. Brand
Department of Sociology
University of California–Los Angeles
264 Haines Hall
Los Angeles, CA 90095-1551
Phone: (310) 266-0826
E-mail: brand@soc.ucla.edu

Abstract

Community colleges are controversial educational institutions, often said to simultaneously expand college opportunities and diminish baccalaureate attainment. We assess the seemingly contradictory functions of community colleges by attending to effect heterogeneity and alternative counterfactual conditions. Using data on postsecondary outcomes of high school graduates of Chicago Public Schools, we find that average effects of community college attendance mask different effects for different students. Enrolling at a community college appears to penalize more-advantaged students who otherwise would have attended four-year colleges; such students represent a relatively small portion of the community college population. However, enrolling at a community college has a modest positive effect on bachelor's degree completion for disadvantaged students who otherwise would not have attended college; these students represent the majority of community college-goers. We conclude that accurately describing the role that community colleges play in social stratification requires analyzing effect heterogeneity and the processes through which heterogeneity arises.

Community colleges are among the most controversial educational institutions (Goldrick-Rab 2010). Scholars alternatively depict them as steering children of the less advantaged away from selective colleges and universities (e.g., Brint and Karabel 1989; Clark 1960; Karabel 1972), or as creating accessible, affordable, and expansionary opportunities for postsecondary education (e.g., Cohen and Brawer 1982; Shaw, Goldrick-Rab, Mazzeo, and Jacobs 2006; Shavit, Arum, and Gamoran 2007). This combination of apparently countervailing effects led Dougherty (1994) to describe the public two-year institution as a "contradictory college." Policymakers are especially concerned about the impact of community college attendance since an increasing number of students attend these colleges but only about one-third obtain any form of credential within six years (Calcagno et al. 2008).

Debates over the function of community colleges typically center on the often analyzed but still poorly understood *average* community college effect, which is interpreted as increasing or decreasing inequality in educational outcomes depending upon whether it is negative or positive, respectively. To the extent that community colleges promote social mobility, individuals who do not attend them are left further behind. But if they diminish opportunities for bachelor's degree completion, gaps in educational attainment are exacerbated. Over the last several decades, dozens of empirical studies have estimated both positive "democratizing" effects of community college as well as negative "diversionary" effects drawing students away from baccalaureate-granting colleges (Alba and Lavin 1981; Alfonso 2006; Anderson 1981; Brint and Karabel 1989; Clark 1960; Doyle 2009; Dougherty 1987, 1994; Dougherty and Kienzl 2006; Goldrick-Rab and Pfeffer 2009; Gonzalez and Hilmer 2006; Grubb 1989, 1991; Hilmer 2000; Leigh and Gill 2003; Long and Kurlaender 2009; Melguizo, Kienzl, and Alfonso 2011; Reynolds [forthcoming]; Reynolds and DesJardins 2009; Rouse 1995, 1998; Sandy, Gonzalez, and Hilmer 2006; Stephan, Rosenbaum, and Person 2009; Velez 1985; Whitaker and Pascarella 1994.). But few researchers have analyzed the degree to which the community college effect is heterogeneous, rather than contradictory, advantaging some students while disadvantaging others.

Given the vast and growing compositional heterogeneity among undergraduates, there is little reason to think that the "treatment" of attending community college affects all students in the same way, as implied by the homogeneity assumption in most empirical estimation approaches. Just as recent work allows for heterogeneity of effects of four-year college completion (Brand and Xie 2010; Brand 2010; Brand and Davis 2011; Brand and Simon-Thomas [forthcoming]; Musick, Brand, and Davis 2012; Xie, Brand, and Jann [forthcoming]), students may respond differently to the community college experience. By attending to community college effect heterogeneity and to the population composition of community college goers (i.e., the characteristics of typical community college students who are more or less likely to attend), as well as to the probable counterfactual paths community college students would have otherwise followed, we can better interpret the average effects of community college attendance. Knowing how the effect of community college attendance varies and the viability of alternative options is also important for the study of social stratification and for educational policy, as more informed advising and targeting of opportunities could reduce inequality in college outcomes and in turn promote life chances.

Using rich longitudinal data on the postsecondary trajectories of students graduating from Chicago Public Schools, we assess heterogeneous treatment effects of community college attendance on bachelor's degree completion using semi- and nonparametric methods based on propensity scores (Xie, Brand, and Jann [forthcoming]). We challenge the implicit homogeneity assumption characteristic of much prior research, and test the possibility that the average democratizing and diversionary effects of community college attendance are in fact different effects for different students. We also consider a range of counterfactual educational paths for individuals who do not attend community colleges, as the relevant alternatives to community college attendance apply to subpopulations of students with different propensities of community college attendance. Although community college students seek many types of degrees and certificates, we focus on the bachelor's degree since it has been the focus of most studies to date and is the college degree that is most indicative of life chances.

We find significant variation in effects on bachelor's degree completion by the likelihood of community college attendance and the probable counterfactual path. High school graduates with a high propensity to attend community college — the majority of the community college population — are most comparable to students who do not attend college, and for them we find a democratizing (positive) effect of community college attendance. Our results indicate that the diversionary (negative) effect of attending community college may be overstated, since it accrues only to a small subpopulation of students who would have otherwise attended a four-year college. We conclude that discussions among education and stratification scholars should move beyond considering the pros and cons of community college attendance for students *in general* to more targeted attention to the implications of community college attendance for *specific* groups of students with differing propensities to attend.

Background and Significance

The community college is a key contributor to the diversity of American higher education (Shavit, Arum, and Gamoran 2007). Public two-year colleges absorbed much of the expansion in postsecondary enrollment that occurred in the mid-20th century and have been repeatedly tasked with maintaining easy access to a college education while also providing a doorway to educational attainment and other socioeconomic opportunities (Cohen and Brawer 1982; Goldrick-Rab, Harris, Mazzeo, Kienzl 2009; Goldrick-Rab 2010). Forty-three percent of all undergraduate students in the U.S. attend community college (American Association of Community Colleges 2011). While often praised for remaining more affordable than other postsecondary options and offering a second chance at educational attainment (Rouse 1995, 1998), the community college has been steadily under attack for low rates of bachelor's degree completion among the community it serves. Some have even suggested that community colleges further socioeconomic disparities in education (Grubb 1991; Brint and Karabel 1989), and that students, especially those seeking bachelor's degrees, are best advised to avoid community college attendance entirely (Guess 2008).

The main explanation for these seemingly disparate judgments is that analysts have focused on different functions of the community college. On the one hand, community colleges exist to provide some college education by opening an affordable, accessible doorway; but on the other hand, they are also expected to facilitate access to baccalaureate granting institutions via transfer. They appear to fulfill the first function fairly well, and the second not as well (Goldrick-Rab 2010). In an attempt to restrict focus to one function of community college, i.e. access to baccalaureate degrees, some analysts restrict their sample to students with the stated intent to attain a bachelor's degree (Alfonso 2006; Doyle 2009; Leigh and Gill 2003; Long and Kurlaender 2009; Whitaker and Pascarella 1994). This approach treats educational expectations, which are known to be malleable and fluctuating (Morgan 2005; Reynolds, Stewart, MacDonald, and Sischo 2006), as static and decisive. Truncating variation among students in studies may also truncate the range of estimated effects of attendance.

But the seemingly incompatible suggestions accorded from past research as to community college effects may hinge upon other empirical and theoretical issues. First, the interpretation of the average effect most commonly estimated by analysts is complicated if there is effect heterogeneity by the probability for community college attendance. Community college attendance may yield positive effects for some subpopulations, and negative for others. Recent research has adopted a propensity score framework to estimate effects (Doyle 2009; Long and Kurlaender 2009; Kalogrides and Grodsky 2011; Melguizo, Kienzl, and Alfonso 2011; Reynolds [forthcoming]; Reynolds and DesJardins 2009; Stephan, Rosenbaum, and Person 2009), stratifying students based on their likelihood of attending community colleges, but that research did not investigate the possibility that the estimated treatment effect differs depending on this likelihood of attendance.[1] Instead, and as in the vast majority of such studies, they estimate average treatment effects and assume away effect heterogeneity. Long and Kurlaender (2009) and Rouse (1995) also use instrumental variable (IV) models to estimate community college effects, where distance to college is the instrument.[2] If there is effect heterogeneity, then IV estimates should be interpreted as local average treatment effects (LATE) that pertain to the population induced to attend community college by the distance, and not to the total population of community college goers. However, neither Long and Kurlaender (2009) nor Rouse (1995) interpret estimated effects as heterogeneous, pertaining to a subpopulation of community college students.

Second, and relatedly, the estimated effect of community college should differ according to the assumed counterfactual educational path, whether it is no immediate postsecondary education or attendance at a non-selective or selective four-year college. This point underscores prior discussions of the divergent functions of the community college, but it is more complex than previously discussed, as it requires precisely identifying how choice sets differ across the population. Community college attendance may *increase* access to educational attainment among the more disadvantaged students, students likely to attend these schools, relative to their most likely counterfactual path – no immediate postsecondary attendance (Grubb 1989; Roderick, Coca, and Nagaoka 2011; Rouse 1995; Sandy, Gonzalez, and Hilmer 2006). That is, if no immediate postsecondary education, rather than four-year college, is the most likely alternative for the majority of community college goers, scholars overstate the penalty to community college attendance by comparing this subpopulation only to four-year college goers. But community college attendance could simultaneously decrease bachelor's degree completion among the more advantaged students, students less likely to attend community college, relative to their probable counterfactual of starting postsecondary education at a four-year college. Among community college-goers, the size of the former population is clearly larger than the size of the latter (though this may be changing over time as more middle-class students are pushed by the recent recession into selecting less expensive college options).

Moreover, most community college students who would instead have attended a four-year college would likely have attended a non-selective four-year school. College selectivity presents disparate opportunities for students, particularly among more disadvantaged students characteristic of community college goers (Alon and Tienda 2005; Brand and Halaby 2006; Dale and Krueger 2011). Thus, studies analyzing community college effects only among college-goers set aside the demonstrably relevant counterfactual of no college attendance (e.g., Doyle 2009; Long and Kurlaender 2009; Reynolds and DesJardins 2009; Whitaker and Pascarella 1994), while others aggregate together institutional types and mask the variable effects of different kinds of colleges.

Analytic Methods

For unit *i*, the effect of community college is defined as the difference between the potential outcomes (in this case, bachelor's degree completion) in the community college state (i.e., the "treated" state, *d*=1) and non-community college state (i.e., the "control" state, *d*=0) (Morgan and Winship 2007; Heckman 2005; Imbens 2004; Rubin 1974):

$$\delta_i = y_i^{d=1} - y_i^{d=0}. \tag{1}$$

Thus, we ask whether students who started at a community college within a year of high school graduation (*d*=1) have different outcomes than they otherwise would have had if they had not begun their post-high school career by enrolling in a community college (*d*=0). It is, of course, impossible to observe both outcomes for the same individual. If unobserved characteristics affect decisions to attend community college and these characteristics are also correlated with eventual bachelor's degree completion, then the estimated effects of community college will be biased. Recent studies of the community college effect have recognized the challenges inherent in establishing causal effects with observational data (Doyle 2009; Long and Kurlaender 2009; Reynolds [forthcoming]; Rouse 1995).

Estimating effects of community college attendance faces a particular causal complexity in how to define the counterfactual state. We decompose the "neutral" or "baseline" counterfactual, that of no community college attendance, by comparing community college attendance to four additional control states: (1) no postsecondary schooling within a year of high school graduation; (2) starting at a four-year nonselective college; (3) starting at a four-year selective college; and (4) starting at a highly selective four-year college. College selectivity is defined by Barron's Profiles of American Colleges 2003, which categorizes colleges according to SAT scores, grade point average, class rank required for admission, and overall admissions acceptance rate. Colleges in the top two categories of Barron's Profiles, ''Most Competitive'' and ''Highly Competitive,'' are considered highly selective for our purposes. The interpretation of the effects corresponding to each of these control states differs; instead of a "baseline" community college effect (in which we compare to no community college attendance), each of these four estimated differences corresponds to the effect of community college attendance relative to the respective alternative.

We begin by estimating simple bivariate associations, or unmatched mean differences, for each of the treatment-control states. We then estimate effects using propensity score matching where individuals are matched according to their propensity for community college attendance relative to the various alternatives (Rosenbaum and Rubin 1983, 1984; Rubin 1997). We estimate propensity scores with a series of probit regressions predicting the propensity of going to community college of the following form:

$$P = p(d_i = 1 \mid X_i) = \Phi(\sum_{k=0}^{K} \beta_k X_{ik}), \quad (2)$$

where P is the propensity score; d_i indicates whether individual i ($i = 1, \ldots, n$) attends community college or not (or one of the four alternative states); and $\boldsymbol{X}$ represents a vector of observed pre-treatment covariates, described in more detail below. As Φ is the cumulative normal distribution, the βs are z scores that indicate the expected change in standard deviation units in the latent dependent variable. These propensity scores represent estimates of individual likelihoods of attending community colleges relative to each control state. The community college effect is the difference in bachelor's degree completion between students with comparable propensities.

We can define treatment effects over several population subsets, but the most commonly investigated parameter using propensity score matching methods is the average treatment effect on the treated (TT):

$$E(\delta \mid d = 1, P) = E(y^{d=1} - y^{d=0} \mid d = 1, P) \quad (3)$$

All matching estimators of the TT take the following general form:

$$TT = \frac{1}{n_1} \sum_{i}^{n_i} (y_{i,d=1} - \sum_{i(j)}^{iJ} w_{i(j)} y_{i(j),d=0}), \quad (4)$$

where n_1 is the number of treatment cases, i is the index over treatment cases, $i(j)$ is the index over untreated cases for treated case i ($i(j)=1,\ldots i(J)$), and $w_{i(j)}$ is the scaled weight (with sum of one) that measures the relative importance of each untreated case. Scholars have not reached a consensus as to which matching estimator performs best in each application (Morgan and Harding 2006), although nearest neighbor and kernel matching, which we use here, have been shown to preform well in simulations (Morgan and Winship 2007). The key to our identification strategy is the assumption that community college attendance is an exogenous event that is not correlated with unobserved factors that could affect bachelor's degree completion. There are at least two advantages to using matching over conventional regression models. First, in contrast to regression, we make no functional form assumption for the relationship between the treatment and outcome using matching. Second, covariate imbalance (also called "common support") is a focal concern in matching routines, while imbalance between treated and untreated cases goes undetected all too often in conventional regression analyses.

While the matching procedure above is an effort to address pre-treatment heterogeneity, we can also extend the propensity score framework to address treatment effect heterogeneity. That is, using semi- and nonparametric methods (Xie, Brand, and Jann [forthcoming]), we consider variation in community college effects by the propensity for community college attendance. The semi-parametric method, called the "stratification-multilevel method" (SM), consists of the following four steps: (1) estimate propensity scores for each unit; (2) construct balanced propensity score strata, such that there are no significant differences in the average values of covariates and the propensity scores between those who do and do not attend community college within strata; (3)

estimate propensity score stratum-specific community college effects (i.e., level-1 effect estimates); and (4) evaluate trends across strata using variance-weighted least squares regressions of the strata-specific effects obtained in step (3) on strata rank at level-2. We estimate our level-2 model by:

$$\delta_s = \hat{\delta}_1 + \gamma S + \varepsilon_s, \quad (5)$$

where level-1 slopes (δ_s) are regressed on propensity score rank indexed by S, $\hat{\delta}_1$ represents the level-2 intercept (i.e., the predicted value of the effect of community college attendance for the lowest propensity stratum), and γ represents the level-2 slope (i.e., the change in the community college effect on bachelor's degree completion with each one unit change to a higher propensity score stratum). Our objective is to look for a systematic pattern of heterogeneous treatment effects across strata. The underlying assumption is that we consider all units within strata, treated and untreated, as homogeneous for estimating treatment effects. Although the assumption of within-stratum homogeneity is unlikely to be true, it is less stringent than the full sample homogeneity assumption. This method has been used in empirical research on the effects of college attendance and completion (Brand 2010; Brand and Davis 2011; Brand and Xie 2010; Musick, Brand, and Davis [forthcoming]; Xie, Brand, and Jann [forthcoming]), but not for community college attendance.

Finally, we test for sensitivity to the parametric and strata-specific homogeneity assumptions imposed in SM using a nonparametric treatment effect heterogeneity method, called the "smoothing-differencing" method (SD), that consists of the following three steps: (1) estimate propensity scores for each unit; (2) for each group (the treatment and control group), fit separate nonparametric regressions of the dependent variable on the propensity score, i.e., local polynomial smoothing (degree 1, bandwidth 0.2); and (3) take the difference in the nonparametric curves between the treated and the untreated to obtain the pattern of treatment effect heterogeneity as a function of the propensity score. This new method has only been used for estimating the effects of college attendance on fertility (Xie, Brand, and Jann [forthcoming]).

Data and Sample

Decisions about attending community colleges are essentially local ones; very few students travel far from home to attend. The same is generally true for students attending non-selective public four-year institutions (Goldrick-Rab 2010; Turley 2009). In large national samples with wide variation among students and colleges, this local concentration can lead to confusion between treatment heterogeneity and treatment *effect* heterogeneity (i.e., different effects of attending community colleges with different characteristics versus different effects of community colleges for different students, respectively). Thus, examining heterogeneous effects of community colleges with national and/or state samples complicates the interpretation of effects relative to considering how attending a specific community college (or set of colleges) exerts heterogeneous effects on the students it aims to serve. Heckman, Ichimura, and Todd (1997) emphasize the importance of comparing treatment and control groups in the same social and economic environment to minimize bias, a consideration that is clearly unmet with national samples. With an eye to addressing these issues, we estimate effects of attending the Chicago City Colleges for the graduates of Chicago Public Schools. The tradeoff is, of course, that we have limited ability to generalize estimates of effects beyond Chicago.

We focus on Chicago because it is among a handful of urban school districts that has for many years had the capacity to follow the trajectories of their graduates, possesses rich data on students' background characteristics and schooling, and represents the nation's fourth largest school district. With funding from the Bill and Melinda Gates Foundation, Chicago worked with the National Student Clearinghouse to implement a low-cost, high-impact approach to identifying the college outcomes of their graduates. Nearly half of students in Chicago Public Schools (CPS) enroll in college within one year of high school, half of those students (24 percent of the sample) enter a college granting bachelor's degrees. More than half of all CPS college-goers enroll in ten in-state colleges, most of them located within Chicago city limits. Among those attending four-year colleges, most are enrolled at schools with graduation rates well below the national norm (Roderick, Coca, and Nagaoka 2011).[3] This is not entirely surprising, since their relatively poor academic qualifications and high level of socioeconomic disadvantage make the set of colleges representing viable choices quite limited (Roderick at al. 2008). The bachelor's degree completion rate within six years of high school is only about 11 percent, low by nearly any standard.

Data on students' pre-college characteristics come from Chicago Public Schools and the uncommonly rich surveys conducted by the Consortium for Chicago School Research (CCSR). We utilize a wide range of measures affecting college choice in Chicago. These include:

1) *Demographic characteristics* (sex, race and ethnicity, citizenship, generational status);

2) *Social background characteristics* (family structure, mothers' education, Census tract social status according to occupation and education, Census tract unemployment and poverty, and Census tract homeowner tenancy);[4]
3) *High school academic achievement* (cumulative grade point average, number of honors courses, number of AP courses, number of absences, placement in special education);
4) *Educational resources* (number of educational resources at home, parental communication, parental involvement);
5) *Educational aspirations and expectations* (college aspirations, college expectations, and parental expectations for college); and
6) *High school characteristics* (percent of students in high school attending a four-year college).

With the notable absence of measures of cognitive ability and direct measures of family income (we have only indirect indicators through neighborhood characteristics), ours is a more extensive set of covariates than typically employed in studies of community college choice. For example, due to limitations in the Ohio administrative source they employed, Long and Kurlaender (2009) can only control on a subset of potentially observable characteristics, and only for a select group of students — those who have taken the ACT and aspire to complete a bachelor's degree. Since the ACT is not required for admission to community college, theirs is an especially selective sample of students that excludes students who wanted to earn a bachelor's degree but did not adequately prepare (a common phenomenon in Chicago; see Roderick et al. 2008) or did not state while in high school that they wanted to earn that degree.

We estimate community college effects for the class of 2001 cohort of CPS graduates. Most analyses estimating community college effects include samples of much older cohorts (Goldrick-Rab 2010). We use a more recent cohort of graduates who nonetheless graduated from high school prior to contemporary reform efforts to increase college attendance in Chicago. We condition our initial sample of CPS graduates (N=14,322) on data availability; specifically students must be included in the National Student Clearinghouse data (N=13,966) and have responded to at least one of the surveys administered by CPS in grades 9 and 11 (N=9,533).[5] Missing information on survey measures is imputed based on all other variables in our models. Departing from some past research, we do not condition the sample on the stated aspiration to earn a bachelor's degree. While some analysts (e.g., Alfonso 2006; Doyle 2009; Leigh and Gill 2003; Long and Kurlaender 2009; Whitaker and Pascarella 1994) restrict the sample to identify comparable control groups, we reject that approach, for the reasons we describe above.[6] Instead, we advance comparability between treated and control groups by examining a local setting, using a rich set of exogenous covariates, restricting our analyses to regions of common support (i.e., no significant covariate or propensity score differences between treated and control groups), and estimating effects to specific subpopulations based on alternative counterfactual conditions.

Results

Descriptive Statistics

College-bound students face many options in Chicago, including more than fifty four-year selective and non-selective colleges and universities, and a system of seven City Colleges (all of which are community colleges). From Table 1, we see that CPS students who attended community college are most similar to individuals who did not attend postsecondary education immediately following college, although some dimensions suggest that students who did not continue their schooling are more disadvantaged than community college-goers. Students who did not continue schooling are less likely to be white, live in more disadvantaged communities during high school, and have lower educational aspirations and expectations. Still, with the exception of race, the differences between community college and non-college goers are smaller than the differences between community college and four-year college goers. Compared to four-year college goers, community college students have less-educated mothers, live more often in non-intact families, reside in substantially more disadvantaged parts of the city (as determined by Census block groups), have lower educational aspirations and expectations, and have lower academic achievement in high school (i.e., attained lower GPAs, taken fewer honors courses and AP classes, and been absent from high school more often). As expected, these differences are largest between community college and highly selective four-year college goers.

College choice is associated with the likelihood of bachelor degree completion for CPS graduates. As described in Table 1, 11.4 percent (N=1,089) of graduates from the CPS class of 2001 earned a bachelor's degree by 2007. This figure includes 1.6 percent (N=74) of students who did not attend any postsecondary school within a year of their high school graduation, 2.8 percent (N=50) of students who started at a community college, 10.7 percent (N=79) of students who attended a nonselective four-year college, 17.5 percent (N=175) of students who attended a selective four-year college, and 54.1 percent (N=711) of students who attended a highly selective four-year college. As we note above, this is a sample marked by high socioeconomic and academic disadvantage, and thus college completion rates are quite low relative to national averages.

[Table 1 About Here]

Propensity Score Probit Regression Estimates

In order to match comparable students to estimate effects of community college attendance on bachelor's degree completion, we first estimate the propensity that a student attends a community college compared to: (1) not attending a community college; (2) not attending postsecondary education within one year of completing high school; (3) attending a non-selective four-year college; (4) attending a selective four-year college; and (5) attending a highly selective four-year college. We fit separate probit regressions for the respective comparisons. The results, reported in Table 2, are as follows:

1) When comparing community college students to those who *did not attend community college* (the all-inclusive, or baseline, comparison), we find that community college students are less likely to have advantaged family backgrounds and be academically prepared for college. The significantly advantaged students who attend highly selective four-year colleges likely drive these differences.
2) When comparing community college students to those who *did not pursue postsecondary schooling* (within a year of high school graduation), we find that high educational expectations predicts community college attendance, but that community college students are less academically prepared for college, net of family background and other factors. As depicted in Table 1, community college students and individuals who do not immediately pursue postsecondary schooling are all disadvantaged. But it may be that community college-goers have the misperception (possibly from

"misaligned ambitions") that they can postpone skill acquisition until postsecondary schooling, whereas students not bound for college may recognize that in their jobs they will have to lean on the skills they gained in high school (Schneider and Stevenson 1999).

3) When comparing community college students to those who *attended a nonselective four-year college*, we find that low educational aspirations, poor academic preparation, and family disadvantage are significant, independent predictors of community college attendance.
4) When comparing community college students to those who *attended a selective four-year college*, the former are significantly disadvantaged with regard to race, family background, high school academic preparation, educational resources, and educational aspirations and expectations of students and parents.
5) When comparing community college students to those who *attended a highly selective four-year college*, we find even larger relative disadvantages as those compared to selective four-year college goers. Community college students are significantly disadvantaged with regard to race, family background, high school academic preparation, educational resources, educational aspirations and expectations of students and parents, and parental involvement.

[Table 2 About Here]

These important differences in observable characteristics, evident between community college students and alternative comparison groups, suggest substantial challenges in estimating overall effects of attending community colleges. Estimating potentially different effects of the consequences of college choice for different subpopulations is our next task.

Propensity Score Matching Estimates

In Table 3, we report propensity score matching results of the treatment effects for the treated (*TT*) under each counterfactual scenario, assuming treatment effect homogeneity. We restrict all analyses to the region of common support (α=0.01). As we should expect, this restriction results in the loss of very few cases for the most appropriate comparisons, i.e. between community college and no immediate college or non-selective four-year college; we lose more cases when we compare community college to selective and highly selective four-year schools. As reported in Table 3, we find that community college-goers are less likely to complete a bachelor's degree compared to individuals not attending community college (our baseline comparison). Unmatched differences suggest a significant 11 percent lower level of bachelor's degree completion (or an 81 percent decrease in the expected odds of completion). Matching reduces our point estimates, suggesting some selection into community college by observed covariates. We do not find notable differences between our three matching methods; nearest neighbor matching with one and five controls and kernel matching estimates all suggest a significant roughly 4 percent lower level of bachelors degree completion (or a 61 percent decrease in the expected odds). This estimate represents an average treatment effect of community college attendance without consideration of effect heterogeneity or the complex set of counterfactual conditions.

In columns 2-4 of Table 3, we provide treatment effects for each of the counterfactual conditions, which allows a much more meaningful interpretation of the community college effect. Compared to students who attend four-year schools, community college students are less likely to complete a bachelor's degree. We find a roughly 3 to 4 percent lower level of bachelor's degree completion (a 56 to 60 percent decrease in expected odds) for community college-goers relative to students starting at a non-selective four-year college when we match on propensity scores. Differences are significant (.05 level) for nearest neighbor matching with five controls and kernel matching, and marginally significant (.10 level) for nearest neighbor matching with one control. We find substantially larger penalties to community college attendance relative to attending a selective four-year college, including a 14 percent lower level (86 percent decrease in expected odds) using single nearest neighbor matching and a roughly 9 percent lower level (79 percent decrease in expected odds) using nearest neighbor matching with five controls and kernel matching. The penalty of community college attendance relative to attending a highly selective four-year college is very large: single nearest neighbor matching suggests a 37 percent lower level (96 percent decrease in the expected odds), and nearest neighbor matching with five controls and kernel matching suggest a 24 and 28 percent lower level, respectively (92 and 94 percent decrease in the expected odds, respectively). Matching estimates of community college attendance relative to selective and highly selective four-year college attendance are all highly significant.

While we find penalties associated with community college attendance compared to attending a four-year college, we also find that community college goers are more likely to obtain a bachelor's degree relative to students who do not immediately pursue postsecondary schooling. Nearest neighbor and kernel matching estimates suggest a 13 percent higher level of bachelor's degree completion (a 93 percent increase in the expected odds). Thus, the overall effect of community college attendance reported in the first column of Table 3 comprises both a large penalty relative to four-year, particularly highly selective, college attendance, in addition to a benefit relative to no immediate postsecondary schooling.

[Table 3 About Here]

Estimates of Effect Heterogeneity

Differences in outcomes reported in Table 3 necessitate a better understanding as to which members of the community college population these educational paths represent viable alternatives. To address this question, we first present descriptive statistics in Table 4 of individual characteristics and educational paths within each stratum for individuals who did and did not attend community college. Strata are constructed such that covariates and the propensity scores are balanced within each stratum. Lower numbered strata correspond to the lower propensity and higher numbered strata correspond to the higher propensity individuals. Among individuals with a high propensity to attend a community college but who did not attend one, the vast majority did not enroll in any college within one year of high school graduation. In other words, for the majority of community college goers, the alternative to community college attendance is not to go to college. For students with a low propensity to attend a community college but who did not attend one, we find proportionately higher levels of four-year college goers. Individuals who have a low propensity for community college attendance have more advantaged social background characteristics, higher academic achievement in high school, more educational resource at home, and higher educational expectations and aspirations than individuals who have a high propensity. These results illustrate the difficulty in interpreting average effects, like those reported in Table 3. The direction and size of the effect depends crucially upon the population upon which its estimation is based, which is influenced by both the location on the propensity score distribution as well as the analyst's choice of the relevant control conditions.

[Table 4 About Here]

Utilizing our balanced propensity score strata, we report estimates of the heterogeneous effects of community college attendance relative to each counterfactual. We begin with the stratification-multilevel method (SM), in which we estimate effects within strata (level-1) and then estimate the trends in effects (level-2).[7] Results are reported in Table 5. We find that the overall effect of attending community college compared to not attending community college on bachelor's degree completion is negative in the low propensity score strata (an 82 percent decrease in the expected odds in stratum 1, $e^{-1.710}$), and positive (albeit not significant) in stratum 6 (a 49 percent increase in the expected odds, $e^{0.397}$). We find a highly significant positive level-2 slope, indicating that the effect of community college increases, and changes direction, as the propensity for community college attendance increases. As we show in Table 4, the appropriate counterfactual varies within the population – in other words, heterogeneity in the treatment effects of community college attendance is related to heterogeneity in the counterfactual. Given the heavy weight in the population composition of the control units towards individuals who did not immediately go to college, the negative effects in strata 2 through 5 are surprising. However, we reason that the comparison to even the relatively small proportion of highly selective four-year college goers in these strata, who are much more likely to obtain bachelor's degrees, are masking the positive effects from the comparison to the relatively large population of individuals who do not immediately pursue college. When we eliminate highly selective four-year college goers from this heterogeneity model, for only one stratum do we find a statistically significant negative effect (results available upon request).

As we note above, heterogeneity in the effects of community college attendance is linked to heterogeneity in the counterfactual. By disaggregating the control state into our four alternative paths, we should thus not expect significant effect heterogeneity. And indeed, we do not find evidence of any remaining significant effect heterogeneity for any of the four counterfactual alternatives. Community college attendance increases the odds of completing a bachelor's degree relative to no immediate postsecondary college in each stratum, with the exception of individuals in stratum 1. Moreover, no clear pattern of individual characteristics differentiates these strata.[8] Attending a community college decreases the odds of completing a bachelor's degree relative to attending a non-selective four-year college in each stratum, significantly in three out of five strata. For this comparison, high propensity students are those who have more disadvantaged family backgrounds and poorer academic achievement in high school. Comparing community college goers to selective four-year college goers, we again observe a penalty for all community college students, with significant negative effects for six out of seven strata. Social background advantages, academic achievement in high school, educational resources and expectations all decrease as the propensity for community college relative to selective four-year college attendance increases. Interestingly, the percentage black also decreases as the propensity for community college increases, while the percentage Hispanic runs in the opposite direction, and the percentage white is relatively stable across strata. Finally, our comparison with highly selective four-year college goers shows very large, highly significant effects for all seven strata. The largest effect in stratum 1 suggests a 95 percent decrease in the expected odds of bachelor's degree completion ($e^{-3.045}$). Like the comparison to moderately selective four-year college attendance, social background advantages, academic achievement in high school, and educational expectations decrease as the propensity for community college relative to highly selective four-year college attendance increases. In this case, however, the percentage white decreases as the propensity for community college increases, while the percentage black runs in the opposite direction, and the percentage Hispanic is relatively stable across strata.

[Table 5 About Here]

To test the sensitivity of our analyses to the linearity assumption we impose in SM, we use the smoothing-differencing heterogeneous treatment effects method (SD), in which we fit separate nonparametric regressions of the dependent variable on the propensity score for the treated and control groups and take the difference in the curves over the propensity score.[9] Results are depicted in Figure 1 for the comparison between community college attendance and no community college attendance ("anything else"), and in Figure 2 (a-d) for the decomposed comparisons with alternative paths. The *x*-axes plot the continuous propensity score and the *y*-axes the differences in nonparametric regressions between treated and controls, i.e. the treatment effect. In Figure 1, we find a significant curvature in the trends in effects, with a negative effect for the low propensity individuals, leveling off to no effect in the middle of the propensity distribution, to slightly positive for the high propensity individuals. Thus, the SD results suggest no effect in the middle of the distribution rather than the negative effect suggested by our results using the SM approach.

[Figure 1 About Here]

We turn our attention to the heterogeneity patterns for the four alternative counterfactuals in Figure 2, in which the SM results suggest no significant linear trend in effects. In Figure 2(a), in which we compare to no immediate college, we see a relatively flat trend in effects, hovering around zero, rather than the positive effects we find using SM. We see a similar trend in effects comparing community college to non-selective and selective four-year college attendance, in Figures 2(b) and 2(c), respectively: a small negative effect that moves towards zero as the propensity for community college attendance, and thus social and academic disadvantage, increases. Figure 2(c) differs from Figure 2(b) in that we see a clear negative effect at the lower tail of the propensity distribution. Finally, in Figure 2(d) we find large negative effects for the comparison with highly selective four-year college attendance, which increase (that is, the negative effect decreases) as the propensity for community college increases. Thus, the students most penalized by attending a community college are those with more advantaged social backgrounds and better academic preparation; it appears these students would be particularly better served by attending a highly selective four-year school.

[Figure 2 About Here]

Discussion and Conclusions

The interpretation of community colleges' role in stratification processes depends on the accurate assessment of the colleges' effects on educational attainment. We have shown that a robust understanding of community college effects requires both an assessment of heterogeneity in effects and a clear specification of the likely alternatives to attending community college for various subpopulations. Using extensive data from Chicago Public Schools, we use propensity score matching to study average effects, and new semi- and nonparametric approaches to study treatment effect heterogeneity. Although examining a single urban context may compromise external validity, it has the advantage of increasing internal validity by reducing unobserved extraneous variation. The Chicago data also offer a rich set of covariates for predicting college-going behavior, useful in countering claims that community college effect estimates suffer from omitted variable bias. We find that attending to effect heterogeneity and the complexity of the appropriate counterfactuals yields a more accurate portrait as to who benefits and who is penalized by attending community college. While enrolling in community college expands the likelihood of bachelor's degree completion for students *who otherwise would not have attended college at all*, it penalizes students *who would otherwise have attended four-year schools*, especially highly selective four-year schools. Moreover, the penalty to community college attendance relative to four-year college attendance is largest among students with a low propensity for community college attendance — students with more advantaged social backgrounds and better academic preparation. This is an important finding given reports of increased enrollment in community colleges among middle-class students, and efforts to boost that enrollment even further.

Yet while it is true that some students would likely be better served by attending four-year rather than two-year colleges, our estimates indicate that relatively few students do this; instead, we find that the most likely alternative to community college attendance for the majority of community college goers is no immediate college attendance at all. It is therefore possible that efforts to focus students' attention on four-year colleges and universities may backfire, effectively reducing educational attainment of those who will choose community colleges for reasons other than the better-expected outcomes associated with four-year institutions.

A few caveats are in order. First, the accuracy of our estimates hinges on whether we have captured all relevant observables that predict community college attendance and bachelor's degree completion. Scholars who have employed an instrumental variables approach [e.g., Long and Kurlaender (2009) and Rouse (1995)] find smaller community college penalties using IV models relative to OLS regression or propensity score models, and suggest this is the result of unobserved heterogeneity. This may be true, but we have offered an alternative explanation for discrepancies in the results across these models: effect heterogeneity. If there is effect heterogeneity, the IV estimates pertain to a subpopulation of students induced to community college by distance to college, who likely do not represent the students deciding between community college and a selective four-year institution, for whom penalties are largest. Nevertheless, the assessment of heterogeneous effects (differences by propensity score strata) could be jeopardized by bias from unobserved selection *if* such selection differs systematically across the distribution of comparable groups and influences degree completion. The bias is conceivably largest at the low end of the propensity score distribution, in which attendance is an unlikely event and thus governed by a host of potentially unobserved factors, and in which we find the largest penalty to community college attendance. Second, we focus on a single urban context, and although doing so potentially increases the internal validity of our results, our findings may nevertheless not generalize to other areas in the United States. The Chicago Public Schools and the City Colleges serve very disadvantaged populations and both are widely known for their challenges with regard to funding and leadership. Third, we examine only one outcome: bachelor's degree completion. Community colleges serve many functions for a diverse population of students. Even those students who we discuss as being "penalized" because they did not complete a four-year degree, may nevertheless have benefited from community college attendance in the labor market, in the marriage market, in their social-psychological well-being and self-acceptance, in their social involvement, and other outcomes that indicate life chances (Hout [forthcoming]; Kalogrides and Grodsky 2011; Kane and Rouse 1995; Marcotte, Bailey, Borkoski, and Kienzl 2005; Rose [forthcoming]).

Having made those cautionary statements, we believe our findings have important implications for policymakers, practitioners, and researchers in higher education. Assuming a homogenous community college "effect" masks variation in effects across the distribution of college goers: We find a benefit to community college attendance among weaker students who would not have otherwise gone to a four-year school, and a penalty to community college attendance among the more advantaged students who would have. Still, the penalty to attending community college should be considered in light of the higher costs associated with attendance at a four-year school. In an era of widespread economic distress facing families alongside rising college tuition costs, we may find an increasing number of students who attend community rather than four-year, even highly selective, colleges. We cannot necessarily predict, however, that this trend will entail a greater number of students facing a penalty of community college attendance, as a shift in the population composition of community college students could coincide with a corresponding shift in its effects.

References

Alba, Richard D. and David E. Lavin. 1981. "Community Colleges and Tracking in Higher Education." *Sociology of Education* 54:223-37.

Alfonso, M. 2006. "The Impact of Community College Attendance on Baccalaureate Attainment." *Research in Higher Education* 47(8):873-903.

Alon, Sigal, and Marta Tienda. 2005. "Assessing the "Mismatch" Hypothesis: Differences in College Graduation Rates by College Selectivity." *Sociology of Education* 78(4):294-315.

American Association of Community Colleges. 2011. "Community College Fast Facts." http://www.aacc.nche.edu/AboutCC/Documents/FactSheet2011.pdf.

Anderson, Kristine L. 1981. "Post-High School Experiences and College Attrition." *Sociology of Education* 54:1-15.

Brand, Jennie E. 2010. "Civic Returns to Higher Education: A Note on Heterogeneous Effects." *Social Forces* 89(2):417-434.

Brand, Jennie E. and Dwight Davis. 2011. "The Impact of College Education on Fertility: Evidence for Heterogeneous Effects." *Demography* 48(3):863-887.

Brand, Jennie E. and Charles N. Halaby. 2006. "Regression and Matching Estimates of the Effects of Elite College Attendance on Educational and Career Achievement." *Social Science Research* 35:749-770.

Brand, Jennie E. and Juli Simon Thomas. "Causal Effect Heterogeneity." [forthcoming] in *Handbook of Causal Analysis for Social Research*, Stephen L. Morgan ed., Springer Series.

Brand, Jennie E. and Yu Xie, 2010. "Who Benefits Most From College? Evidence for Negative Selection in Heterogeneous Economic Returns to Higher Education." *American Sociological Review* 75(2):273-302.

Brint, Steven, and Jerome Karabel. 1989. *The Diverted Dream: Community Colleges and the Promise of Educational Opportunity in America*, *1900-1985*. Oxford: Oxford University Press.

Calcagno, Juan Carlos, Thomas Bailey, Davis Jenkins, Gregory Kienzl, and Timothy Leinbach. 2008. "Community College Student Success: What Institutional Characteristics Make a Difference?" *Economic of Education Review* 27:632-645.

Carneiro, Pedro, James J. Heckman, and Edward Vytlacil. 2011. "Estimating Marginal Returns to Education." *American Economic Review* 101(6):2754-2781.

Clark, Burton A. 1960. "The 'Cooling-out' Function in Higher Education." *American Journal of Sociology* 65(6):569–576.

Cohen, Arthur M. and Florence B. Brawer. 1982. *The American Community College*. San Francisco: Jossey-Bass Publishers.

Dale, Stacey and Alan B. Krueger. 2011. "Estimating the Return to College Selectivity over the Career Using Administrative Earnings Data." National Bureau of Economic Research Working Paper No. 17159.

Dougherty, Kevin J. 1987. "The Effects of Community Colleges: Aid or Hindrance to Socioeconomic Attainment?" *Sociology of Education* 60(2):86-103.

Dougherty, Kevin J. 1994. *The Contradictory College: The Conflicting Origins, Impacts, and Futures of the Community College*. Albany, NY: State University of New York Press.

Dougherty, Kevin J., and Gregory S. Kienzl. 2006. "It's Not Enough to Get Through the Open Door: Inequalities by Social Background in Transfer from Community College to Four-Year Colleges." *Teachers College Record* 108(3):452-487.

Doyle, William R. 2009. "The Effect of Community College Enrollment on Bachelor's Degree Completion." *Economics of Education Review* 28:199-206.

Goldrick-Rab, Sara. 2010. "Challenges and Opportunities for Improving Community College Student Success." *Review of Educational Research* 80(3):437-469.

Goldrick-Rab, Sara, Douglass N. Harris, Christopher Mazzeo, and Gregory Kienzl. 2009. "Transforming America's Community Colleges: A Federal Policy Proposal to Expand Opportunity and Promote Economic Prosperity." *Metropolitan Policy Program at Brookings.*

Goldrick-Rab, Sara and Fabian Pfeffer. 2009. "Beyond Access: Explaining Social Class Differences in College Student Mobility." *Sociology of Education* 82(2): 101-125

Gonzalez, Arturo, and Michael J. Hilmer. 2006. "The Role of Two-Year Colleges in the Improving Situation of Hispanic Postsecondary Education." *Economics of Education Review* 25(3):249-257.

Grubb, W. Norton. 1989. "The Effects of Differentiation on Educational Attainment: The Case of Community Colleges." *Review of Higher Education* 12:349-74.

Grubb, W. Norton. 1991. "The Decline of Community College Transfer Rates: Evidence from National Longitudinal Surveys." *Journal of Higher Education* 62(2):194-222.

Guess, Andy. 2008. "A Penalty for Starting at a Community College?" *Inside Higher Ed.* October 1, http://www.insidehighered.com/news/2008/10/01/pathway.

Heckman, James J. 2005. "The Scientific Model of Causality." Pp. 1–98 in *Sociological Methodology,* Vol. 35, edited by Ross M. Stolzenberg. Boston, MA: Blackwell Publishing.

Heckman, James J., Hidehiko Ichimura, and Petra E. Todd. 1997. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program." *The Review of Economic Studies* 64(4):605–654.

Hilmer, Michael J. 2000. "Does the Return to University Quality Differ for Transfer Students and Direct Attendees?" *Economics of Education Review* 19(1):47-61.

Hout, Michael. [forthcoming]. "Social and Economic Returns to College Education." *Annual Review of Sociology*.

Imbens, Guido W. 2004. "Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review." *The Review of Economics and Statistics* 86:4–29.

Jann, Ben, Jennie E. Brand, and Yu Xie. 2010. "HTE: Stata module to perform heterogeneous treatment effect analysis by propensity score stratification with a multilevel model." in Stata: ssc install hte; http://ideas.repec.org/c/boc/bocode/s457129.html;

Kalogrides, Demetra and Eric Grodsky. 2011. "Something to Fall Back On: Community Colleges as a Safety Net." *Social Forces* 89(3):853-878.

Kane, Thomas J., and Cecilia Elena Rouse. 1995. "Labor Market Returns to Tow- and Four-Year Colleges." *American Economic Review* 85(3):600-614.

Karabel, Jerome. 1972. "Community Colleges and Social Stratification." *Harvard Educational Review* 42:251–62.

Leigh, D. E., and A. M. Gill. 2003. “Do Community Colleges Really Divert Students from earning Bachelor’s Degrees?” *Economics of Education Review* 22(1):23–30.

Long, B. T. and Kurlaender, Michal. 2009. “Do Community Colleges Provide a Viable Pathway to a Baccalaureate Degree?” *Educational Evaluation and Policy Analysis* 31(1):30-53.

Marcotte David E., Thomas Bailey, Carey Borkoski, Greg S. Kienzl .2005. “The Returns of a Community College Education: Evidence from the National Education Longitudinal Survey (NELS).” *Educational Evaluation and Policy Analysis* 27:157-175.

Melguizo, T., G. Kienzl, and M. Alfonso. 2011. “Comparing the Educational Attainment of Community College Transfer Students and Four-Year Rising Juniors using Propensity Score Matching Methods.” *Journal of Higher Education* 82(3):265-291.

Morgan, Stephen L. 2005. *On the Edge of Commitment: Educational Attainment and Race in the United States*, Stanford University Press.

Morgan, Stephen, and David Harding. 2006. “Matching Estimators of Causal Effects: Prospects and Pitfalls in Theory and Practice.” *Sociological Methods and Research* 35(1):3-60.

Morgan, Stephen, and Christopher Winship. 2007. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge, UK: Cambridge University Press.

Musick, Kelly, Jennie E. Brand, and Dwight Davis. 2012. "Variation in the Relationship Between Education and Marriage: Marriage Market Mismatch?" *Journal of Marriage and Family* 74:53-69.

Reynolds, C. Lockwood. [forthcoming]. "Where to Attend? Estimating the Effects of Beginning College at a Two-Year Institution." *Economics of Education Review.*

Reynolds, C. Lockwood, and Stephen L. DesJardins. 2009. "The Use of Matching Methods in Higher Education Research: Answering Whether Attendance at a 2-Year Institution Results in Differences in Educational Attainment." *Higher Education: Handbook of Theory and Research* 24:47-97.

Reynolds, John, Michael Stewart, Ryan MacDonald, and Lacey Sischo. 2006. "Have Adolescents Become Too Ambitious? High School Seniors' Educational Plans, 1976-2000." *Social Problems* 53(2):186-206.

Roderick, Melissa, Vanessa Coca, and Jenny Nagaoka. 2011. "Potholes on the Road to College: High School Effects in Shaping Urban Students' Participation in College Application, Four-Year College Enrollment, and College Match." *Sociology of Education* 84(3):178-211.

Rose, Mike. [forthcoming]. *Back to School: Second Chances at Higher Education.* New York: The New Press.

Rosenbaum, Paul R., and Donald B. Rubin. 1983. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 70(1):41–55.

Rosenbaum, Paul R., and Donald B. Rubin. 1984. "Reducing Bias in Observational Studies Using Subclassification on the Propensity Score." *Journal of the American Statistical Association* 79:516–24.

Rouse, Cecilia E. 1995. "Democratization or Diversion—The Effect of Community-Colleges on Educational Attainment." *Journal of Business and Economic Statistics* 13(2):217–224.

Rouse, Cecilia E. 1998. "Do Two-Year Colleges Increase Overall Educational Attainment? Evidence from the States." *Journal of Policy Analysis and Management* 17(4):595–620.

Rubin, Donald. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66:685-701.

Rubin, Donald. 1997. "Estimating Causal Effects from Large Data Sets Using Propensity Scores." *Annals of Internal Medicine* 127:757–63.

Sandy, Jonathon, Arturo Gonzalez, and Michael Hilmer. 2006. "Alternative Paths to College Completion: Effect of Attending a Two-Year School on the Probability of Completing a Four-Year Degree." *Economics of Education Review* 25(5):463-471.

Schneider, Barbara and David Stevenson. 1999. *The Ambitious Generation: America's Teenagers Motivated by Directionless*. New Haven: Yale University Press.

Shavit, Yoshi, Richard Arum, and Adam Gamoran (eds.). 2007. *Stratification in Higher Education: A Comparative Study*. Palo Alto, CA:Stanford University Press.

Shaw, Kathleen, Sara Goldrick-Rab, Christopher Mazzeo, and Jerry Jacobs. 2006. *Putting Poor People to Work: How the Work-First Idea Eroded College Access for the Poor*. New York: Russell Sage.

Stephan, Jennifer L., James E. Rosenbaum and Ann E. Person. 2009. "Stratification in College Entry and Completion." *Social Science Research* 38(3): 572–593.

Turley, Ruth N. López. 2009. "College Proximity: Mapping Access to Opportunity." *Sociology of Education* 82(2):126-146.

Velez, William. 1985. "Finishing College: The Effects of College Type." *Sociology of Education* 58:191-200.

Whitaker, David G., and Ernest T. Pascarella. 1994. "Two-year College Attendance and Socioeconomic Attainment: Some Additional Evidence. *Journal of Higher Education* 65(2):194-210.

Xie, Yu. 2011. "Population Heterogeneity and Causal Inference." *University of Michigan Population Studies Center Research Report* 11-731.

Xie, Yu, Jennie E. Brand, and Ben Jann. [forthcoming]. "Estimating Heterogeneous Treatment Effects with Observational Data." *Sociological Methodology*.

[1] Prior work has tested for variation between treatment and selected covariates. For example Long and Kurlaender (2009) find some evidence of variation in effects by gender, race, and achievement level (ACT score).

[2] Distance to college is an invalid instrument if it affects four-year degree completion directly, rather than only indirectly through community college attendance. In fact, we know that distance to four-year colleges increases the likelihood of attending four-year colleges, as it is frequently used as an instrument in studies of the effects of four-year college attendance (e.g., Carneiro, Heckman, and Vytlacil 2011).

[3] The most popular community colleges for CPS students are City Colleges: Wilbur Wright, Richard J. Daley, and Harold Washington. The most popular four-year colleges are the University of Illinois at Chicago (highly selective), Northeastern Illinois University (non-selective), the University of Illinois at Urbana-Champaign (highly selective), Chicago State University (selective), Northern Illinois University (selective), Columbia College of Chicago (non-selective), and the Southern Illinois University at Carbondale (selective).

[4] The neighborhood measures have been constructed by CCSR as follows: Neighborhood social status is the standardized mean of the percentage of persons 16 years or older who are managers and executives in a Census block and the (logged) mean level of education among people over 18 years. Neighborhood unemployment and poverty is based on the percent of males over 18 years old employed one or more weeks during the year and the percent of families above the poverty line in a Census block. Neighborhood homeownership is the average number of years of tenancy of homeowners in the census block.

[5] We also excluded a small number of students who attended a private two-year college. Although public attention given to private two-year colleges warrants analyses of these students, in our sample their number is too low to allow separate analysis (46 out of 14,332 CPS 2001 graduates, i.e. 0.3 percent, started at a private two-year institution or a less than two-year institution). And we excluded students attending a four-year institution of unknown selectivity (247 cases, i.e. 1.7 percent). Given the low share of these combined cases (2.2 percent of CPS 2001 graduates) and the absence of strong arguments why

this particular kind of missingness should be related to our outcome of interest, we consider it unlikely that this exclusion significantly biases our estimates.

[6] Other studies have gone even further to try and identify "true" comparison groups. For example, Melguizo, Kienzl, and Alfonso (2011) compare the college outcomes of community college transfer students and four-year college rising juniors. It is perhaps unsurprising that they find no negative community college effect, since theirs is instead the effect of being a community college transfer student — a selective, high achievement, group of community college attendees.

[7] For each within-stratum regression, we adjust for unbalanced covariates, i.e. those covariates that significantly differ from one another within-strata. We use the Stata module -hte- for SM (Jann, Brand, and Xie 2010).

[8] Results analogous to Table 4 for the alternative counterfactual comparisons are available upon request.

[9] We use the Stata module -hte2- for SD developed by Ben Jann, Jennie E. Brand, and Yu Xie (available upon request from these authors).

Table 1

Descriptive Statistics by Educational Path: Chicago Public Schools Sample (n=9,533)

	Full Sample	No Immediate College	Community College	Non-Selective 4-Yr College	Selective 4-Yr College	Highly Selective 4-Yr College
Demographic Characteristics						
Female	0.591	0.553	0.609	0.622	0.657	0.638
	[0.492]	[0.497]	[0.488]	[0.485]	[0.475]	[0.481]
White	0.126	0.089	0.131	0.164	0.123	0.231
	[0.331]	[0.285]	[0.337]	[0.370]	[0.328]	[0.421]
Black	0.519	0.528	0.538	0.468	0.731	0.329
	[0.500]	[0.499]	[0.499]	[0.499]	[0.444]	[0.470]
Hispanic	0.297	0.353	0.275	0.278	0.131	0.259
	[0.457]	[0.478]	[0.447]	[0.449]	[0.337]	[0.438]
Other race	0.059	0.030	0.056	0.091	0.016	0.182
	[0.235]	[0.170]	[0.230]	[0.287]	[0.125]	[0.386]
U.S. born	0.801	0.811	0.812	0.769	0.914	0.680
	[0.399]	[0.391]	[0.391]	[0.422]	[0.280]	[0.467]
Second generation	0.390	0.378	0.403	0.405	0.178	0.569
	[0.488]	[0.485]	[0.491]	[0.491]	[0.383]	[0.495]
Social Background Characteristics						
Intact family	0.247	0.236	0.244	0.249	0.240	0.297
	[0.431]	[0.424]	[0.429]	[0.433]	[0.427]	[0.457]
Mother college graduate	0.124	0.098	0.113	0.120	0.192	0.181
	[0.329]	[0.297]	[0.317]	[0.325]	[0.394]	[0.385]
Neighborhood social status	-0.235	-0.357	-0.266	-0.073	-0.024	-0.010
	[0.800]	[0.758]	[0.760]	[0.789]	[0.790]	[0.909]
Neighborhood non-poor	0.225	0.319	0.218	0.105	0.272	-0.071
	[0.799]	[0.764]	[0.827]	[0.829]	[0.820]	[0.771]
Neighborhood homeowner	11.399	11.146	11.510	11.125	12.749	11.280
	[4.198]	[4.034]	[4.238]	[4.130]	[4.735]	[4.126]
High School Academic Achievement						
Cumulative GPA	2.493	2.220	2.221	2.645	2.789	3.528
	[0.821]	[0.715]	[0.653]	[0.687]	[0.660]	[0.611]
Honors courses	0.791	0.458	0.380	0.835	1.188	2.207
	[1.330]	[1.025]	[0.909]	[1.320]	[1.486]	[1.596]
AP credits	0.145	0.079	0.042	0.103	0.145	0.538
	[0.495]	[0.364]	[0.252]	[0.407]	[0.468]	[0.879]
Absences	6.352	7.488	6.948	5.351	4.862	3.184
	[6.407]	[7.008]	[6.588]	[5.164]	[4.602]	[3.869]
Special education	0.083	0.116	0.095	0.068	0.020	0.005
	[0.276]	[0.320]	[0.294]	[0.251]	[0.140]	[0.0728]
Educational Resources						
No. of educ. resources at home	1.018	0.943	0.945	1.019	1.305	1.164
	[1.041]	[0.980]	[0.995]	[1.058]	[1.145]	[1.163]
Parental communication	2.664	2.577	2.667	2.639	3.116	2.637
	[2.734]	[2.679]	[2.737]	[2.792]	[2.784]	[2.822]
Parental involvement	3.932	3.908	3.964	3.830	4.451	3.639
	[3.987]	[3.969]	[4.017]	[4.036]	[3.991]	[3.945]
Aspirations and Expectations						
College aspirations	0.237	0.173	0.218	0.268	0.371	0.374
	[0.426]	[0.379]	[0.413]	[0.443]	[0.483]	[0.484]
College expectations	0.197	0.135	0.178	0.208	0.328	0.339
	[0.398]	[0.341]	[0.383]	[0.406]	[0.470]	[0.473]
College expectations parents	0.865	0.841	0.861	0.850	0.924	0.917
	[0.342]	[0.366]	[0.346]	[0.357]	[0.265]	[0.276]
High School Characteristics						
College-going rate	0.519	0.541	0.534	0.507	0.494	0.447
	[0.100]	[0.0918]	[0.0926]	[0.0976]	[0.100]	[0.100]
Outcome						
Bachelor's degree completion	0.114	0.016	0.028	0.107	0.174	0.541
	[0.318]	[0.124]	[0.166]	[0.309]	[0.380]	[0.498]
N	9,533	4,704	1,772	740	1,003	1,314

Notes: Numbers in brackets are standard deviations. Sample restricted to CPS high school graduates with college transcript data.

Table 2

Probit Regression Estimates for Models Predicting Community College Attendance

	Community College Attendance vs.				
	Anything Else	No Immediate College	Non-Selective 4-Yr College	Selective 4-Yr College	Highly Selective 4-Yr College
Female	0.146 ***	0.184 ***	0.085	0.158 **	0.078
	(0.032)	(0.036)	(0.059)	(0.061)	(0.072)
Black	-0.127 *	-0.109	0.073	-0.564 ***	-0.543 ***
	(0.062)	(0.070)	(0.109)	(0.111)	(0.128)
Hispanic	-0.385 ***	-0.476 ***	-0.218 *	-0.197 †	-0.421 ***
	(0.058)	(0.065)	(0.100)	(0.116)	(0.115)
Other race	0.007	0.124	-0.208	0.355 †	-0.426 **
	(0.082)	(0.101)	(0.131)	(0.188)	(0.137)
U.S. born	0.187 ***	0.212 ***	0.163 †	-0.039	0.052
	(0.050)	(0.055)	(0.092)	(0.109)	(0.100)
Second generation	0.284 ***	0.299 ***	0.363 ***	0.551 ***	-0.005
	(0.052)	(0.057)	(0.095)	(0.103)	(0.110)
Intact family	0.023	0.012	0.053	0.032	0.155
	(0.044)	(0.049)	(0.081)	(0.080)	(0.095)
Mother college graduate	0.008	0.034	0.056	-0.079	-0.099
	(0.053)	(0.060)	(0.095)	(0.088)	(0.109)
Neighborhood social status	-0.059 *	0.004	-0.249 ***	-0.160 ***	-0.118 *
	(0.024)	(0.027)	(0.045)	(0.047)	(0.051)
Neighborhood non-poor	-0.100 ***	-0.112 ***	-0.104 *	-0.079 †	-0.074
	(0.026)	(0.029)	(0.047)	(0.048)	(0.056)
Neighborhood homeowner	0.006	0.009 *	0.016 *	-0.005	0.008
	(0.004)	(0.004)	(0.007)	(0.007)	(0.009)
HS GPA	-0.244 ***	-0.028	-0.521 ***	-0.741 ***	-1.382 ***
	(0.028)	(0.031)	(0.052)	(0.053)	(0.068)
HS honors courses	-0.119 ***	-0.059 **	-0.066 *	-0.132 ***	-0.171 ***
	(0.017)	(0.021)	(0.029)	(0.027)	(0.027)
HS AP credits	-0.207 ***	-0.251 ***	-0.001	-0.032	-0.173 *
	(0.051)	(0.063)	(0.090)	(0.086)	(0.072)
HS absences	-0.005 †	-0.007 **	0.006	0.012 *	0.007
	(0.003)	(0.003)	(0.005)	(0.005)	(0.007)
HS special education	-0.053	-0.118 *	0.041	0.692 ***	0.676 **
	(0.054)	(0.057)	(0.104)	(0.141)	(0.231)
Educ. resources	-0.063 **	-0.062 *	-0.042	-0.103 **	-0.082 †
	(0.022)	(0.025)	(0.039)	(0.039)	(0.047)
Parental communication	0.013	0.015	-0.020	0.018	-0.018
	(0.015)	(0.016)	(0.028)	(0.029)	(0.034)
Parental involvement	0.000	-0.009	0.028	0.027	0.069 **
	(0.010)	(0.011)	(0.020)	(0.020)	(0.024)
College aspirations	0.027	0.098	-0.209 *	-0.176 †	-0.361 **
	(0.056)	(0.062)	(0.099)	(0.096)	(0.116)
College expectations	0.023	0.142 *	0.084	-0.175 †	-0.239 *
	(0.059)	(0.066)	(0.104)	(0.099)	(0.116)
College expectations (parental)	0.029	0.079	0.092	-0.246 **	-0.299 **
	(0.046)	(0.050)	(0.082)	(0.095)	(0.110)
HS college-going rate	0.512 **	-0.387 †	1.166 ***	2.710 ***	3.497 ***
	(0.176)	(0.199)	(0.316)	(0.307)	(0.369)
Constant	-0.689 ***	-0.549 ***	0.614 *	1.266 ***	3.073 ***
	(0.147)	(0.165)	(0.257)	(0.264)	(0.323)
LR χ^2	471.1	230.9	306.9	977.9	2344.8
P > χ^2	0.000	0.000	0.000	0.000	0.000
N	9,533	6,476	2,512	2,775	3,086

Notes: Numbers in parentheses are standard errors. Sample restricted to CPS high school graduates with college transcript data.
† p<.10 * p < .05 ** p < .01 *** p < .001 (two-tailed tests)

Table 3

Matching Estimates of Community College Attendance on Bachelor's Degree Completion

	Community College Attendance vs.				
	Anything Else	No Immediate College	Non-Selective 4-Yr College	Selective 4-Yr College	Highly Selective 4-Yr College
Unmatched Differences	-0.106 ***	0.012 **	-0.079 ***	-0.146 ***	-0.513 ***
	(0.008)	(0.004)	(0.010)	(0.010)	(0.013)
Nearest Neighbor Matching (*k*=1)	-0.041 ***	0.010 †	-0.034 †	-0.138 ***	-0.371 ***
	(0.008)	(0.006)	(0.020)	(0.033)	(0.109)
Nearest Neighbor Matching (*k*=5)	-0.044 ***	0.013 **	-0.034 *	-0.091 ***	-0.244 **
	(0.006)	(0.005)	(0.017)	(0.028)	(0.087)
Kernel Matching	-0.043 ***	0.013 **	-0.040 *	-0.084 ***	-0.278 ***
	(0.006)	(0.004)	(0.016)	(0.024)	(0.068)
N	9,533	6,476	2,512	2,775	3,086
N (On Common Support)	9,532	6,468	2,512	2,747	3,014

Notes: Numbers in parentheses are standard errors. Propensity scores were estimated by probit regression models of community college attendance on the set of pre-college covariates as described in Table 2. Sample restricted to CPS high school graduates with college transcript data.

† p<.10 * p <.05 ** p < .01 *** p < .001 (two-tailed tests)

Table 4

Descriptive Statistics by Strata and Treatment Condition: Community College vs. Anything Else

	Stratum 1		Stratum 2		Stratum 3		Stratum 4		Stratum 5		Stratum 6	
	d=0	*d*=1	*d*=0	*d*=1	*d*=0	*d*=1	*d*=0	*d*=1	*d*=0	*d*=1	*d*=0	*d*=1
Treatment												
No Immediate College	0.261	0.000	0.501	0.000	0.644	0.000	0.750	0.000	0.817	0.000	0.888	0.000
Community College	0.000	1.000	0.000	1.000	0.000	1.000	0.000	1.000	0.000	1.000	0.000	1.000
Non-Selective 4-Year	0.083	0.000	0.120	0.000	0.101	0.000	0.096	0.000	0.089	0.000	0.077	0.000
Selective 4-Year	0.172	0.000	0.172	0.000	0.153	0.000	0.109	0.000	0.063	0.000	0.022	0.000
Highly Selective 4-Year	0.485	0.000	0.207	0.000	0.102	0.000	0.045	0.000	0.031	0.000	0.012	0.000
Demographic Characteristics												
Female	0.581	0.533	0.528	0.577	0.548	0.543	0.603	0.578	0.644	0.711	0.730	0.683
White	0.192	0.156	0.073	0.115	0.076	0.067	0.090	0.103	0.136	0.139	0.302	0.284
Black	0.390	0.344	0.485	0.390	0.573	0.568	0.591	0.627	0.571	0.584	0.400	0.416
Hispanic	0.298	0.433	0.389	0.473	0.326	0.341	0.277	0.219	0.239	0.229	0.215	0.152
Other race	0.120	0.067	0.053	0.022	0.025	0.023	0.042	0.052	0.054	0.048	0.083	0.148
U.S. born	0.747	0.722	0.767	0.797	0.806	0.809	0.837	0.834	0.841	0.844	0.787	0.770
Second generation	0.440	0.456	0.359	0.390	0.324	0.326	0.357	0.342	0.404	0.391	0.609	0.667
Social Background Characteristics												
Intact family	0.290	0.256	0.240	0.269	0.215	0.207	0.237	0.230	0.244	0.255	0.280	0.292
Mother college graduate	0.188	0.133	0.133	0.126	0.102	0.085	0.110	0.110	0.108	0.125	0.073	0.128
Neighborhood social status	-0.001	-0.330	-0.254	-0.305	-0.298	-0.319	-0.297	-0.235	-0.290	-0.249	-0.312	-0.218
Neighborhood non-poor	0.034	0.172	0.290	0.226	0.376	0.391	0.324	0.317	0.171	0.114	-0.111	-0.108
Neighborhood homeowner	11.370	10.980	10.820	10.830	11.220	11.050	11.580	11.480	11.890	12.080	11.570	12.180
High School Academic Achievement												
Cumulative GPA	3.583	3.218	2.892	2.867	2.469	2.447	2.123	2.118	1.827	1.879	1.585	1.727
Honors courses	2.956	2.949	1.060	1.000	0.344	0.345	0.140	0.131	0.047	0.052	0.032	0.026
AP credits	0.626	0.490	0.074	0.122	0.023	0.019	0.010	0.000	0.003	0.003	0.004	0.000
Absences	3.740	5.552	5.054	6.270	6.453	6.783	7.380	7.142	8.216	7.325	8.235	7.277
Special education	0.008	0.022	0.062	0.088	0.110	0.106	0.101	0.110	0.116	0.108	0.103	0.062
Educational Resources												
No. of educ. resources at home	1.306	1.089	1.079	1.198	1.030	0.948	0.958	0.932	0.813	0.841	0.690	0.872
Parental communication	2.801	2.389	2.590	2.868	2.565	2.457	2.703	2.563	2.599	2.530	2.655	3.374
Parental involvement	3.958	3.644	3.840	4.231	3.819	3.615	4.041	3.913	3.945	3.799	3.919	4.790
Aspirations and Expectations												
College aspirations	0.363	0.233	0.235	0.231	0.217	0.183	0.207	0.195	0.194	0.224	0.146	0.296
College expectations	0.322	0.156	0.187	0.176	0.186	0.171	0.160	0.161	0.156	0.167	0.110	0.255
College expectations (parental)	0.901	0.844	0.866	0.857	0.857	0.829	0.848	0.867	0.859	0.875	0.858	0.889
High School Characteristics												
College-going rate	0.454	0.490	0.498	0.511	0.520	0.527	0.541	0.534	0.560	0.554	0.571	0.552
Outcome												
Bachelor's degree completion	0.371	0.067	0.159	0.022	0.083	0.036	0.041	0.039	0.021	0.003	0.012	0.021
N	1,465	90	1,201	182	1,791	387	1,766	517	941	353	493	243

Notes: Sample restricted to CPS high school graduates with college transcript data.

Table 5

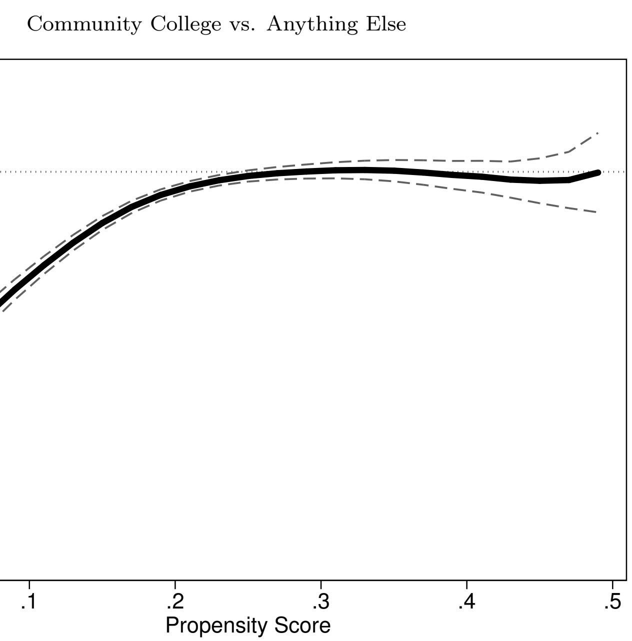
Stratification Multilevel (SM) Estimates of Community College Attendance on Bachelor's Degree Completion

	Community College Attendance vs.				
	Anything Else	No Immediate College	Non-Selective 4-Yr College	Selective 4-Yr College	Highly Selective 4-Yr College
Level-1					
Stratum 1	-1.710 ***	-0.179	-0.672	-1.210 *	-3.045 ***
	(0.436)	(0.627)	(0.540)	(0.554)	(0.741)
Stratum 2	-2.130 ***	0.552 †	-0.996 **	-1.726 **	-2.025 ***
	(0.512)	(0.299)	(0.342)	(0.610)	(0.511)
Stratum 3	-0.883 **	0.946 *	-1.159 **	-1.375 ***	-2.015 ***
	(0.285)	(0.442)	(0.412)	(0.390)	(0.423)
Stratum 4	-0.069	1.209 *	-0.050	-1.268 **	-2.811 ***
	(0.257)	(0.607)	(0.487)	(0.460)	(0.534)
Stratum 5	-2.034 *	0.336	-2.013 **	-2.151 ***	-2.593 ***
	(1.027)	(0.521)	(0.652)	(0.449)	(0.445)
Stratum 6	0.397			-0.184	-2.389 ***
	(0.633)			(0.781)	(0.484)
Stratum 7				-2.051 **	-3.027 ***
				(0.744)	(0.564)
Level-2	0.500 ***	0.113	-0.075	-0.059	-0.092
	(0.125)	(0.164)	(0.171)	(0.118)	(0.107)
N	9,533	6,476	2,512	2,775	3,086

Notes: Logit coefficients reported; numbers in parentheses are standard errors. Propensity scores were estimated by probit regression models of community college attendance on the set of pre-college covariates as described in Table 2. Sample restricted to CPS high school graduates with college transcript data.

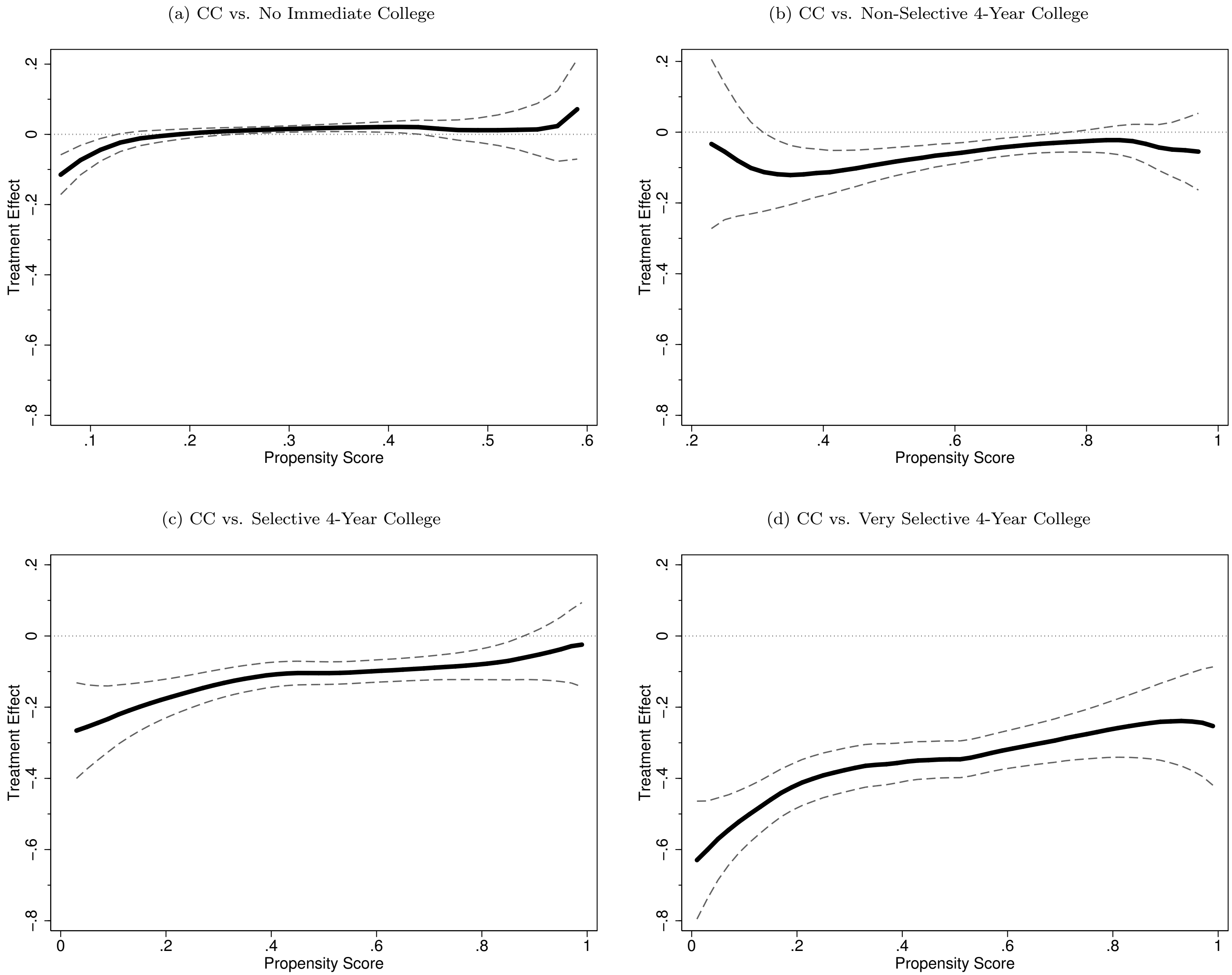
† p<.10 * p <.05 ** p < .01 *** p < .001 (two-tailed tests)

Figure 1: Smoothing-Differencing (SD) Method for Heterogeneous Treatment Effects



Note: Solid line indicates local polynomial fit. Dashed lines indicate 95 percent confidence interval

Figure 2: Smoothing-Differencing (SD) Method for Heterogeneous Treatment Effects



Note: Solid line indicates local polynomial fit. Dashed lines indicate 95 percent confidence interval