NBER WORKING PAPER SERIES

DO COMMUNITY COLLEGES PROVIDE A VIABLE PATHWAY TO A BACCALAUREATE DEGREE?

Bridget Terry Long
Michal Kurlaender

Working Paper 14367
http://www.nber.org/papers/w14367

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2008

The authors thank the Ohio Board of Regents and in particular Rod Chu, Darrell Glenn, Robert Sheehan, and Andy Lechler for help with the data. Eric Bettinger provided invaluable support during the project. All opinions and mistakes are our own. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Do Community Colleges provide a Viable Pathway to a Baccalaureate Degree?
Bridget Terry Long and Michal Kurlaender
NBER Working Paper No. 14367
September 2008
JEL No. C1,I2,J24

ABSTRACT

Community colleges have become an important entryway for students intending to complete a baccalaureate degree. However, many question the viability of the transfer function and wonder whether students suffer a penalty for starting at a two-year institution. This paper examines how the outcomes of community college entrants compare to similar students who initially entered four-year institutions within the Ohio public higher education system. Using a detailed dataset, we track outcomes for nine years and employ multiple strategies to deal with selection issues: propensity score matching and instrumental variables. The results suggest that straightforward estimates are significantly biased, but even after accounting for selection, students who initially begin at a community college were 14.5 percent less likely to complete a bachelor's degree within nine years.

Bridget Terry Long
Harvard University
Graduate School of Education
Gutman Library 465
6 Appian Way
Cambridge, MA 02138
and NBER
longbr@gse.harvard.edu

Michal Kurlaender
University of California, Davis
One Shields Avenue
2051 Academic Surge
Davis, CA 95616
mkurlaender@ucdavis.edu

I. INTRODUCTION

Community colleges play an important role in American higher education. For many, they offer open, affordable access to postsecondary schooling. Nearly one-half of all undergraduates at public institutions attend a community college, and low-income, minority, and first-generation students are more likely to attend community colleges than four-year institutions.[1] While community colleges serve a diverse set of needs, their success is often measured by their ability to facilitate students' transfer to four-year institutions, and in recent years, community colleges have become an increasingly important pathway to a baccalaureate degree. Capacity constraints at many four-year institutions have prompted states to rely on the community college sector to accommodate much of the continued expansion of higher education (Evelyn, 2002).[2] In addition, rising admissions requirements and escalating tuition costs at four-year institutions have limited access to these schools for some students (Mills, 2006).

The viability of the community college transfer function has long been a source of debate. On the one side, the "democratization" view suggests that community colleges have increased overall access to postsecondary education by making higher education possible for many students with open admissions and low costs (Rouse, 1995; Rouse, 1998; Leigh & Gill, 2003). On the other hand, some posit that community colleges serve as a "diversion" by channeling students into vocational courses and weakening their educational progress with increased flexibility and nontraditional patterns of attendance (e.g. delayed entry, part-time enrollment, combining employment with schooling) (Grubb, 1991; Brint & Karabel, 1989; Dougherty, 1994). In reality, many community college students face a number of personal and structural barriers to transfer, including financial concerns, limited information, and the lack of coordination between two-year and four-year institutions. As such, estimates from the most

[1] Source: U.S. Department of Education, National Center for Education Statistics, IPEDS enrollment data and *The Condition of Education 2000.*

[2] As noted by Evelyn (2002), some states are even considering using community colleges to award bachelor's degrees in areas of specific need to the state.

recent national longitudinal postsecondary study reveal that only 37 percent of students who graduated high school in 1992 and began at a community college eventually transferred to a four-year college (Adelman 2005).

Although previous research has highlighted the reduced likelihood of baccalaureate attainment for students who begin at a community college, the bulk of research in this area focuses on students who graduated high school over fifteen years ago and even thirty years ago.[3] Given the recent trends highlighted above, a more timely analysis is needed to reexamine and extend our understanding of the current function of community colleges.[4] This paper provides a new perspective on the role of community colleges for a cohort of students who entered Ohio public colleges and universities in fall 1998.

The paper makes several important contributions to the existing literature on community colleges. First, we use a detailed, robust data source that is an improvement over the data of many previous studies. We track students over a longer period of time than any other study to our knowledge (nine years) and across multiple institutions to capture transfer behavior and periods of stopping out and returning to college. Also, because our dataset is a complete census of everyone in the public higher education system in Ohio, attrition in our sample is small.[5]

---

[3] Studies using the National Longitudinal Survey of Youth (NLSY79) (e.g, Leigh & Gill, 1997; 2003) focus on students who were ages 14-22 when first interviewed in 1979 suggesting they graduated high school in the 1970s or 1980s. Studies using the High School & Beyond (e.g., Rouse, 1995) focus on students who graduated high school in 1982. Studies using the National Education Longitudinal Study (NELS88) (e.g., Roksa, 2006; Alfonso, 2006) focus on students who graduated high school in 1992.

[4] Other recent developments that have likely affected the role of community colleges include increases in high school graduation requirements, which may have improved the academic preparation of some students while influencing others to drop out and seek their GEDs at community colleges, and the increasing use of loans beginning with the 1992 Higher Education Act Reauthorization, which may have affected student choice between two- and four-year college options. At the same time, many community colleges opted out of giving students federal student loans due to government rules that penalize colleges with high default rates.

[5] Regardless of where a student enrolls or how many times they change institutions, they are tracked within the system using their unique Social Security number. This is not the case, however, in many other studies that have used national data sets. Although these sources are designed to be nationally representative, they suffer from attrition over time, especially in the later follow-ups.[5] For example, the fourth follow-up study of NELS had a target of 15,649 students, but only 12,144 are considered

Given that all community college students do not aim to complete a baccalaureate degree, one must also account for degree intent in this sort of analysis (Alfonso, 2006; Alfonso, Bailey, & Scott, 2005). Our dataset includes important information on student degree intent, and so we are able to disentangle the many populations that community colleges serve to focus on students with a four-year degree objective. Finally, unlike many other state unit record datasets, we have information on family background to control for other differences between students.[6]

Research on the transfer function of community colleges must address a number of empirical difficulties, which have been undertaken with varying degrees of success in past literature. Students do not randomly choose their colleges, and the types of students who first attend a community college differ from those who initially enroll in a four-year institution. To address this selection, past studies have employed a variety of strategies: Heckman two-stage sample selection model (Alfonso, 2006), propensity score matching (Stephan & Rosenbaum, 2006; Melguizo, Kienzl, & Alfonso, 2007) or instrumental variables (Rouse, 1995). Each method has its own strengths and weaknesses, and so, unlike any other study, this paper utilizes

---

completed cases resulting in a response rate of 77.6 percent (NCES, 2002). Adjusting the sample weights in response to this attrition may not be an adequate solution due to the fact that there are likely to be important, unobservable differences between those in the sample and those missing. Therefore, information from these databases on behavior several years out of high school may not truly be nationally representative.

[6] Most state administrative databases lack significant detail on community college students because information on academic preparation and test scores is not required for admission and financial aid records are often not part of the data systems. (Even when included, not all students complete the aid application and so are missing from the dataset.) Therefore, many studies based on administrative data only have basic demographic information (e.g., Moore & Shulock, 2007; Leigh & Gill, 2007). Efforts to supplement the data with student surveys often have low response rates. For example, the CUNY administrative data utilized by Bailey and Weininger (2002) was supplemented with a survey on students' social background and place of birth, but the response rate was only 31.4 percent (sample weights had to be developed for the analysis). An exception is the Florida administrative data used by studies such as Calcagno, Crosta, Bailey, and Jenkins (2007). As a K-20 Data Warehouse, it also includes information on high school preparation, test scores, and family income (if a FAFSA was submitted). At the other end of the spectrum, the SUNY data used by Ehrenberg and Smith (2004) is grouped so that analysis cannot be done at the individual level. In contrast, the data used in this paper has degree intent information from the college application and details about family income, high school preparation and achievement, and high school type from the student surveys that accompany college entrance exams (i.e., the ACT and SAT).

two different methodologies (propensity scores and instrumental variables) to triangulate the impact of starting at a community college on baccalaureate attainment. This also allows us to explore the relative effectiveness of each method in dealing with selection. The propensity scoring method gets its strength from the wealth of observable information we have on student characteristics and preparation. In contrast, the instrumental variables strategy accounts for unobservable differences by using, among other things, the importance of college proximity in college choice. Distance is a particularly relevant issue in the choices of Ohio students, and so our use of distance as an instrument to deal with selection is more convincing than in other contexts. The analysis examines the postsecondary outcomes of students who initially enter higher education through the community college system in comparison to those who began at a four-year institution.

The propensity score matching and instrumental variables results point to similar conclusions. Each suggests that straightforward regression estimates are biased, and so simple comparisons between two-year and four-year students, which suggest students who initially enroll at a community college do far worse, should be treated with caution. Additionally, unobservable differences between students appear to be important, thereby suggesting that the instrumental variable estimates do a better job reducing bias as a result of self-selection. However, even after accounting for selection, it appears that students who begin at community colleges still suffer a penalty. According to the instrumental variable estimates, the most conservative of our models, community college students were 14.5 percent less likely to complete a bachelor's degree in nine years in comparison to similar students who began at a four-year university. The paper is organized as follows: in Section II we discuss relevant literature addressing these questions, Section III describes our empirical strategies, Section IV details our results, and Section V concludes.

II. LITERATURE REVIEW

The standard human capital model, developed by Becker (1993), provides some insight for why students might choose to attend a community college over four-year institutions. First, in an effort to minimize costs, students may select to attend community colleges because they are less expensive than four-year institutions and can reduce the total costs of a bachelor's degree. Community colleges also give students time to improve their educational record, and therefore, increase their likelihood of being admitted to a higher-quality, degree-granting, four-year institution. Hilmer (1997) finds that students choose higher quality universities if they first attend community colleges, with the largest quality increases being observed for students who come from poor families or who perform poorly in high school.

Partly because community college participants vary widely in their educational backgrounds and purposes of participation, it is difficult to characterize the performance of community colleges in the educational attainment process. Although community colleges were originally designed to serve as "junior colleges" to four-year programs, today they are an incredibly diverse set of institutions with varied missions and courses of study (Alfonso, Bailey, & Scott, 2005). In addition to serving traditional-age students who recently graduated from high school, community colleges enroll many older, non-traditional students who may take classes for personal enrichment rather than attempting to complete a degree. Moreover, community colleges attract students who are often "experimenting" with postsecondary education, without necessarily the drive to complete a degree (Grubb, 1991). Nevertheless, community colleges' success is often measured in their ability to facilitate students' transfer to four-year institutions and complete a bachelor's degree.

Many question whether community colleges offer a genuine path to attaining a baccalaureate degree. As noted above, the increased flexibility and nontraditional patterns of attendance afforded by community colleges may weaken a student's progression through postsecondary education and make transferring to a four-year college less likely (Grubb, 1991; Brint & Karabel, 1989; Dougherty, 1994). Furthermore, the administrative costs of transferring to a four-year college may be exceedingly prohibitive for many students (Anderson, Sun, & Alfonso, 2006). Some also argue that the educational system is fraught with market failures as a result of lack of information about the process for postsecondary entry, (e.g. financial aid procedures and admissions standards at four-year institutions), and other barriers that may limit the opportunity sets of students from disadvantaged backgrounds (Hilmer, 1998).

On the other hand, a "democratization" view suggests that community colleges have increased overall access to postsecondary education with open admissions and low cost. In fact, Rouse (1995), utilizing the High School and Beyond (HSB) dataset, finds that starting at a two-year college does not affect the likelihood of attaining a bachelor's degree for those students intending to transfer to a four-year college. Although she does estimate a negative "diversion" effect, the "democratization" effect of increased access to higher education is estimated to be larger. Using a slightly different methodology, Leigh and Gill (2003), controlling for a student's desired level of education, find similar results to Rouse using the National Longitudinal Survey of Youth (NLSY79).

There has also been a substantial amount of research describing who utilizes the transfer route from community colleges. Utilizing the National Educational Longitudinal Study (NELS), Adelman (2005) reports that among 1992 twelfth-graders whose first postsecondary institution was a community college, 37 percent transferred to a four-year college, and 60 percent of those that transferred had earned a bachelor's degree by 2000. Using administrative data on a cohort of students initially enrolled in California community colleges in 1999-2000, Moore and Shulock (2007) report that among the 60 percent of degree seekers attending California community colleges, about 25 percent transfer to a university and/or earn an associate's degree within six years. Overall, several papers have concluded that those who transfer from a community college to a four-year college are of a higher social class, have higher academic preparation, are less likely to be minority, and less likely to be female than are those who do not transfer (Dougherty & Kienzl, 2006; Lee & Frank, 1990; Dougherty, 1987, 1994; Whitaker & Pascarella 1994; Grubb, 1991; Surette, 2001).

Questions also exist about whether community college students experience the same returns to education as their counterparts who entered four-year colleges initially. Kane and Rouse (1995) use transcript data from the National Longitudinal Study of the Class of 1972 (NLS72) and the NLSY79 to compare the returns to a college credit at two-year and four-year colleges. Controlling for family income and measured ability, they find wage differentials for both two-year and four-year college credits are positive and of similar magnitude. In separate work, Leigh and Gill (1997) (also using the NLSY79), examine whether returns to a community college education differ for adults who return to college as compared to continuing high school graduates. They find positive returns of essentially the same magnitude. Finally, Hilmer (2000) examines whether the return to university quality differs for community college transfer students and direct attendees. Using the HSB data, he finds that the return to university quality differs dramatically across both educational paths, and institutional quality ranges with the length of time spent at the initial institution having a significant negative effect on transfer students.

Researchers have also debated whether there is a return to getting an education at a community college without completing a degree. Kane and Rouse (1995) found a positive return to attending a two-year college even when a degree was not obtained. Grubb (1995) also finds similar results after corrections to his NLS72 data; in a revision of a 1993 paper where he concludes that economic benefits may accrue to even modest levels of community college participation.

While these studies have laid the foundation on the impact of community colleges on baccalaureate attainment and labor market outcomes, this study provides an important comprehensive update, using data on students age 17 to 20 who entered Ohio public colleges and universities in Fall 1998. The data include important details such as educational intent and family background and tracks students nine years across public institutions within the state. In addition, by taking a state-level perspective, the paper provides insight that is difficult to discern from studies using national data. States differ incredibly in how they choose to use community colleges within their statewide higher education systems. In some states, community colleges are the primary access points (e.g., California), while other states focus all remediation there (e.g., Florida), and still others favor four-year college entry with the community colleges primarily serving less traditional needs such as older students and workforce retraining efforts (e.g. Massachusetts). Because of this state-level variation, it can be difficult to fully understand community college dynamics when looking at national data. For instance, the marginal community college student (i.e., the student on the fence between a four- and two-year college) differs by state, and this makes it hard to interpret some of the patterns observed in national data. Our interpretation might differ greatly if looking at a community college student in California versus one in Massachusetts due to the fact that their backgrounds, preparation, and aspirations are likely to be different as a result of their contrasting state higher education structures. By using state-level data, we are able to be much more specific about who the marginal student is and take into account the state context in our interpretation.

This study also sheds light on a previously-unstudied, important context: the fifth largest public higher education system in the country. While the external validity of our results are limited by focusing on a particular state, the higher education system, population demographics, and diverse labor markets within Ohio are similar to larger national trends and many neighboring states (Mortenson, 2002), and so our results could be used to reflect on issues facing many other states. It is worth noting however that because our study and many others on this topic focus on traditional-age college students, we can say little about the many older students who attend community colleges.

III. EMPIRICAL FRAMEWORK AND METHODOLOGY

Data

This paper uses a longitudinal, administrative dataset of students in the Ohio public higher education system. Maintained by the Ohio Board of Regents (OBR), the data include information from several sources including applications, college transcripts, and entrance exams and their accompanying questionnaires. Therefore, we have detailed information about each student's family background, high school preparation, postsecondary intentions, college performance and course-taking behavior, and degree completion.[7] In addition, because the data cover the entire public higher education system in Ohio, we are able to track students across schools and determine the outcomes of individuals who may have continued their educations or completed their degrees at different postsecondary institutions from the ones they originally entered. With these data, we examine several student outcomes including baccalaureate degree completion (after four, six, and nine years), total credit hours completed, and stopout behavior (within one, two, or six years).[8] One drawback to the data is that it does not include students who transfer to private colleges or transfer out of the state; these students may incorrectly be designated as dropouts. However, this potential measurement error is likely to be very small

[7] We use three main files submitted by all public colleges and universities in Ohio to the OBR Higher Education Information (HEI) System: Student Entrance, Student Enrollment, and Course Enrollment. In addition, we use test score and student background information that is the result of an OBR data match with ACT, Inc. All of these files, which also denote the institution, are linked by the student's Social Security number (we receive an anonymous version of the data), and a student may have files from multiple institutions if they attended more than one school during the time period. These data are available for each term (fall, spring, summer, and for those on the quarter system, winter) from fall 1997 to spring 2007.

[8] In our paper, students who end college attendance sometime during the nine-year period, do not return to any statewide public institution, and have not completed a bachelor's degree are considered stop/dropouts. The multiple measures of stopout behavior denote the last time the student was enrolled in college.

because, according to analysis of Integrated Postsecondary Education Data System (IPEDS) data, the percentage of students thought to transfer to such schools is a small fraction of the total number of observed dropouts.[9]

This study focuses on first-time freshmen age 17 to 20 who began their postsecondary study during fall 1998 and follows them for nine years until spring 2007. Due to the focus on baccalaureate completion, we restrict the sample to individuals who have demonstrated the intent to complete four-year degree either on their college application or by starting at a four-year university. OBR collects information on the educational intent of all community college students on their applications with eight possible options, and we define baccalaureate degree intent as having designated the intention either: (i) "To obtain a bachelor's degree;" (ii) "To obtain an associate degree for transfer;" or (iii) "To transfer before completing a degree or certificate."[10] Responses to this question about educational intent are completed privately and in no way affect outcomes such as admissions or access to certain classes, and so we believe the variable is a fairly accurate depiction of educational goals at the time of college entry. In fall 1998, 60 percent of first-time community college students indicated intent to obtain a bachelor's degree or transfer to a four-year college with or without an Associate Degree.[11]

---

[9] While it is the case that we can not track students who transfer to private institutions or public out-of-state institutions, this is not likely to be a large group. The IPEDS, a NCES annual database of postsecondary institutions in the United States, tracks the number of transfers at each institution but does not record the state of residence of transfer students. However, it does track the states of residence for incoming freshmen. Assuming that the distribution of transfer students matches the geographical distribution of the incoming freshmen class, then one would expect around 650 Ohio students to transfer to non-public or non-Ohio schools. If we further assume that *all* 650 transfer students just finished their first year of school, then about 4.3 percent of observed dropouts in our data are actually transfer students.

[10] The other options were: (iv) "To obtain knowledge for personal interest;" (v) "To upgrade skills for current job by taking only selected courses;" (vi) "To train for a new career by taking only selected courses;" (vii) "To obtain a certificate;" and (viii) "To obtain an associate degree for the job market." Students who began at a four-year institution are not asked about intent and are assumed to want a baccalaureate degree.

[11] The percentage at community colleges intending to get a bachelor's degree is slightly smaller in our data than that found in other surveys (e.g., NELS and the Beginning Postsecondary Study). Beyond slight differences in the way the question is posed to students in each survey, one reason for the different percentages may be related to *how* the questions are asked. There are concerns that students in other

We also restrict the sample to students who took the ACT.[12] This requirement further emphasizes that the sample, particularly the community college students, had some baccalaureate degree intent as most four-year colleges require students to take either the ACT or SAT, though most schools do not require a minimum score. The ACT restriction is also necessary because this source gives information on student achievement and other important controls for the models. Because of this restriction, our results may reflect a more conservative estimate of the penalty felt by community college students as we may be inadvertently excluding some students who do want a baccalaureate degree but failed to prepare for admission into a four-year university. Finally, due to the use of proximity as an instrument, we restrict the sample to Ohio residents (i.e., distance would not serve as a good instrument for Florida residents attending an Ohio college). Beyond these sample requirements, we drop observations without complete college transcript or zip code information.[13]

Although this paper focuses on students in Ohio, the results have external validity. The percentage of Ohio public school students who graduate from high school and enter college the following fall are near the national averages (Mortenson, 2002), and the higher education system, demographics, and diverse labor markets within Ohio are similar to many neighboring states and national conditions. Moreover, Ohio plays a prominent role in higher education as it has the fifth largest public higher education system in the United States (after California, Florida, New York,

---

surveys report higher educational aspirations, which may not match the realities surrounding college admission and/or enrollment (i.e. academic preparation, financial constraints, etc. (Rosenbaum, 2001). The low-stakes method in which degree intent is collected in our dataset could mean that our survey yields more honest answers.

[12] Approximately two-thirds of high school students in Ohio take the ACT exam. Our dataset includes the highest score of the student and his or her most recent responses to the ACT survey, which includes self-reported information on high school preparation and performance as well as the intended plan of study in college.

[13] After applying the basic sample restrictions (age 17-20; four-year degree intent; took the ACT; Ohio resident), we had to drop 1,092 observations (3.1 percent of the sample) due to missing complete college transcript information and 3,517 observations (10 percent of the sample) due to missing zip code information. While the number missing zip code information is large, it is mainly driven by the fact that Kent State University did not submit any zip code information resulting in us dropping its students from the sample (2,396 observations or 68 of those missing zip code information).

and Texas) (NCES, 2000). While our results could be used to reflect on community college issues in some other states, caution should be taken when making comparisons to states with very different postsecondary systems, such as California.

Basic Empirical Strategy

Our research question asks how the outcomes of community college entrants compare to similar four-year college students. We address this question using regression techniques, beginning with a basic Ordinary Least Squares (OLS) model estimating the impact of beginning at a community college on the total number of college credits completed:

(1) $CREDITS_i = \alpha_{1i} + \beta_1\ COMMCOLL_i + \beta_2\ X_{1i} + \varepsilon_{1i}$

where $COMMCOLL_i$ is a dummy variable equal to one if the student initially began at a community college (zero otherwise), and β_1 is the parameter associated with beginning at a two-year college. X_i is a vector of background controls including student demographics (e.g., race, gender, age, and age squared), family background (e.g., parents' income and income squared), and student ability (e.g., percentile ACT score and the score squared).[14] The error term is represented by ε_i. For the outcomes variables that are dichotomous in nature (e.g., degree completion within nine years), we employ probit models instead of OLS.

Our regression models control for observable differences between students in the sample to isolate the impact of beginning at a community college. However, other factors may influence who chooses to initially attend a community college versus four-year institution. Stated more formally, our probit and OLS estimates may lead to biased estimates if ε_{1i} is correlated with college choice due to unmeasured factors. For this reason we employ two additional empirical

---

[14] To maintain the largest sample size possible, students missing parental income information were included by created a "missing parental income" dummy variable. This had to be done for 1,316 observations, or 8.9 percent of the sample. However, for this reason, the coefficients on parental income should not be used in interpretation.

strategies to deal with selection. The first is to use propensity scores to match similar students on observable characteristics, and the second uses an instrumental variable approach.

Addressing Selection using a Propensity Matching Approach

Students' choices about whether to enroll in two-year versus four-year colleges are subject to their preferences, financial constraints, academic profile, beliefs about the prospects of benefiting from the respective institutions, and other unobservable characteristics. To address the bias that results from this potential selection, we first use a propensity score blocking technique to model selection into two-year colleges versus nonselective four-year institutions. Propensity scores are useful for establishing a comparison group that is similar, on average, to a treatment group (in this case two-year colleges) along a set of observed characteristics. First, we assembled a set of variables to predict choice of two-year versus four-year postsecondary entry and estimate the following binomial logistic regression model:

(2) $\Pr(\text{COMMCOLL})_i = \alpha_{3i} + \delta X_{3i} + \varepsilon_{3i}$

where $\Pr(\text{COMMCOLL})_i$ is the probability for the i^{th} student to enter postsecondary study as a function of variety demographic characteristics and academic performance indicators, X_i (summary information on the estimation of the propensity score is included in Appendix Table 1). Thus, the propensity score is a "single number that indicates the extent to which one person is similar to another along a collection of observed characteristics" (Agodini & Dynarski, 2004). A histogram of the estimated propensity score for the two-year entrants versus four-year starters is in Appendix Figure 1.

Next, we use the propensity score (fitted probabilities of two-year selection from this model) to stratify the sample into an optimal number of subclasses or "blocks." In stratifying the sample, we sort from the lowest to highest estimated propensity scores, discarding the observations with an estimated propensity score less than the minimum or greater than the maximum estimated propensity score for treated units (Dehejia & Wahba, 1999). The goal is to stratify the sample into groups of respondents with similar observed chances of postsecondary entry via two-year versus four-year institutions. The strata are chosen so that the covariates are "balanced" within each stratum, meaning that there are no statistically significant differences in means of covariates between those who started at two-year and those who started at four-year institutions (Dehejia & Wahba, 1999).[15] We present mean propensity scores for two-year entrants (treated) and four-year entrants (control) by strata in Appendix Table 2 We then conduct our subsequent substantive models testing the effect of starting at a community college as fixed effects models on the propensity score blocks, thereby evaluating the within-stratum difference in the outcomes between two-year and four-year entrants (Dehejia & Whaba, 1999).[16] Although the approach does not solve the selection problem on unobservables, it is an effective way to control for observed heterogeneity in type of postsecondary entry (Rubin, 1997; Morgan, 2001; Rosenbaum & Rubin, 1983).

Addressing Selection using an Instrumental Variables Approach

Our second strategy to address the issue of selection is an instrumental variables (IV) approach. In order to be a valid instrument, one must identify a variable that influences the choice of a two-year versus four-year institution while not impacting the outcomes of interest (i.e., persistence, degree completion, and stop out behavior). For this paper, we use proximity as instruments. Previous research has shown that students are more likely to attend one school over another based on how close the colleges are to their homes (Rouse, 1995; Long, 2004). More specifically, our instruments are the distance from the student's home to the closest two-year

[15] To ensure that the individual covariates are balanced within each stratum, we include additional higher order terms of the background variables, see Appendix, Table 1.

[16] For a more detailed discussion of propensity score "binning" procedures see Dehejia and Wahba (2002); Conniffe, Gash, and O'Connell (2000); Dehejia and Wahba (1999); Rosenbaum and Rubin (1983). Several studies demonstrate that propensity score impacts are not sensitive to the ways in which matching is conducted (Dehejia & Wahba, 2002; Smith & Todd, 2000)

college and distance to the closest nonselective four-year university.[17] The IV approach is described as follows:

(3) $CREDITS_i = \alpha_{4i} + \theta_1 COMMCOLL_i + \theta_2 X_{i4i} + \varepsilon_{4i}$

where: $COMMCOLL_i = f(DIST\_2YR_i, DIST\_NONSEL4YR_i) + \nu_i$

As in the earlier equations, $CREDITS_i$ is the outcome variable, and $COMMCOLL_i$ is a dummy variable equal to one if the person began at a two-year college. The difference with this approach is that $COMMCOLL_i$ is instrumented for in the second part of the equation with the distance from the student's home to the closest two-year college and nonselective four-year university, represented by $DIST\_2YR_i$ and $DIST\_NONSEL4YR_i$, respectively.

There have been several critiques of using distance as an instrument. First, Rouse (1995) and Card (1995) caution that the estimates should be interpreted with caution because families who value education might choose to live in close proximity to colleges, and therefore distance could be endogenous. If so, the instruments would be invalid because residence would also be correlated with educational outcomes (due to family preferences). Additionally, the relevance of distance is problematic when using national data sets because of differences in the way distance is perceived across the country and extreme variation in the geographical dispersion of colleges across a state.[18] However, these concerns are limited in this context. First, in Ohio, the colleges and universities are purposely distributed throughout the state so that each resident is within 30 miles of a school. During the 1960s and 1970s, with the goal of improving college access in Ohio, Governor James Rhodes influenced the location of colleges so that every state resident was located within 30 miles of a college campus (OBR, 2001). Therefore, the four-year and two-year

[17] The distances were calculated from the student's high school to the nearest of each type of postsecondary institution. In Ohio, the nonselective, four-year universities are open admissions schools.

[18] These critiques have focused on the power of proximity as a predictor. Clearly, distance might be perceived differently in Nevada versus New Jersey. Also, the geography of college locations differs greatly by state as this creates very different norms in college attendance patterns by distance. However, this paper focuses on one state in which students have a strong pattern of attending the closest college to their home.

colleges are scattered throughout the state, and each resident is close to one or both of each type of college regardless of their underlying preferences. Proximity is a particularly relevant factor in Ohio enrollment patterns. The median distance from a student's home to the college of attendance is only 26 miles in our sample with nearly 60 percent of students attending a college within 50 miles of their homes. In addition, due to the fact that we focus on traditional-age college students, their residential decisions are likely due to parental decisions made when the student was in high school or even earlier, long before the college choice decision (Rouse, 1995).

Results from the first stage of the IV are displayed in Appendix Table 3.[19] In addition to the instruments, the regression also includes the following exogenous variables: a dummy variable for female students (baseline: male students), dummy variables for Black, Hispanic, Asian, and Native American students (baseline: White students), age, age-squared, parental income, parental income-squared, a dummy indicating missing values for parental income, and the student's math and English ACT scores. The signs of the estimates of the first stage of IV are as expected: our sample of students was more likely to choose to initially attend a community college the closer the nearest two-year college and the farther away the nearest four-year nonselective university. The test of joint significance for the instruments suggests that they have substantial explanatory power in the first stage with an F statistic of 3263.11.

IV. RESULTS

Descriptive Analysis: Comparing students with Four-year Degree Intent?

Table 1 displays summary statistics for our sample. The sample focuses on students who began their postsecondary study at age 17-20 years old, who not only indicate four-year degree intent but also took the ACT, an additional sign of aspirations to get a baccalaureate degree. As

[19] These first-stage results apply to the IV estimates discussed below in Table 3, Panel C, and Table 4, specifications (3) and (6).

discussed above, previous empirical work suggest that community college students enter postsecondary schooling with lower academic credentials on average than their counterparts at four-year institutions (as measured by high school grade point average and ACT scores). However, the differences are not large between students at nonselective four-year universities in Ohio and those at community colleges. Thus, for most of our analysis we focus more explicitly on students at nonselective, four-year institutions as the main comparison group to community college students.[20]

The bottom part of Table 1 displays the mean enrollment and educational outcomes of students at each type of college. Students at each type of institution experience very different outcomes. Among all students who start at two-year colleges, nearly 44 percent have dropped out or stopped out six years after starting, as compared with 34 percent who start at nonselective, four-year institutions and 18 percent who start at selective, four-year institutions. The rates of departure are consistently higher for those who start at community colleges versus four-year institutions, regardless of timing of exit. More importantly, it is hard to ignore the vast differences in bachelor's degree receipt between those who start at two-year colleges and those who begin at four-year institutions. Among community college students *with the demonstrated intention to get a four-year degree*, only 26 percent obtain a bachelor's degree within nine years of starting. Meanwhile nearly twice and three times as many students who begin at nonselective (50 percent) and selective four-year institutions (73 percent), do so. However, these comparisons do not account for differences between the students at each type of institution and should not be

[20] It is worth noting that our sample characteristics are different than those of college students generally due to the fact that we limit the sample by age and to those who indicate wanting a four-year degree and who took the ACT. In the general population, one would find much starker differences between students at selective four-year colleges and community colleges. Meanwhile, the data suggest that the student bodies at nonselective four-year colleges and community colleges have much more similar backgrounds and academic preparation even without our sample restrictions. Students at nonselective four-year and two-year (i.e. community colleges) institutions tend to be older, are more likely to be non-White, and have lower academic performance levels. Students at these schools are also more likely to attend part-time or less-than-part-time.

interpreted as causal effects. Community college students are also less likely to be enrolled full-time, with 56 percent of two-year students enrolling full-time as first-time freshmen, as compared to 82 percent of students enrolled at selective, four-year institutions and 67 percent of students enrolled at nonselective, four-year institutions. This may be a decision endogenous to college choice as community colleges offer more flexibility to be able to attend part-time.

Is there a Community College Penalty? Results from the Regression Analysis

Table 2 displays the relative likelihood that students who began at different types of institutions complete a degree within six or nine years of initial enrollment. The six-year benchmark, or 150 percent of the time needed to complete a four-year degree full-time, is commonly used in the literature to calculate graduation rates, but we also show nine-year graduation rates due to the fact that many community college students enroll in higher education part-time. The table reports marginal effects estimated by probit regression models of starting at a community college. However, linear probability models produce similar results.[21]

The estimates suggest that students who start at a community college versus a four-year institution, in particular, in comparison to a selective four-year university, are significantly less likely to have completed a baccalaureate during the time frame. Specifications 1 and 6 show the simple comparison without any controls in the model. As shown in the upper part of the table, community college students were 43 percent less likely to complete a bachelor's degree than a student who began at a selective four-year college. However, given differences between students who begin at community colleges versus four-year schools, the following specifications control for student demographic characteristics (race, gender, age), parents' income, and student ability (ACT scores). Even after accounting for these differences, although lower in magnitude (from -

[21] For instance, specification (9) with all of the controls yields a probit estimate of -0.2023 (0.0078). Using OLS instead, the estimate is -0.1913 (0.0074). Likewise, specification (10) yields a probit estimate of -0.2174 (0.0081). Using OLS instead, the estimate is -0.2046 (0.0076).

0.426 to -0.363), the estimated penalty nevertheless persists. There is also an estimated penalty, albeit a smaller one, for students who attend nonselective four-year institutions in comparison to their counterparts at selective four-year institutions. The F-test comparing the results of attending a community college versus a nonselective four-year university suggests that the estimates are statistically different from each other.

The bottom part of the table limits the sample only to students at nonselective institutions, both four-year and two-year. As discussed above, these students are much more similar across institutions. Again, we note the lower likelihood of degree completion among those who start at a community college. We also note that the estimated penalty is smaller now that we have omitted the comparison with selective four-year institutions. In comparison to the results in the upper part of the table, however, there is only a modest change in the size of the estimate when including controls for background and ability (from -0.225 to -0.202). This suggests that the sample of students at nonselective schools is already very similar and so the covariates do not explain much of the variation in outcomes (below we focus on making comparisons between two-year and nonselective four-year entrants). The lower likelihood of graduation for community college entrants remains larger even after extending the time frame to nine years. Nevertheless, selection into two-year versus four-year colleges may still be influenced by a variety of factors not accounted for in our model, thus next we turn to alternate specifications that address potential selectivity bias.

Comparing the Probit to the Propensity Score and IV Results

Table 3 presents a series of fitted models investigating the effect of starting at a two-year college versus a nonselective four-year institution on a host of outcomes — bachelor's receipt within nine years (column 1), six years (column 2), and four years (column 3), and total credit hours earned after six years (column 4). In addition, the last column looks at the negative outcome of stopping out within six years without completing a degree (column 5). Panel A presents the results without any correction for selection, Panel B presents the propensity score specification results, and Panel C presents the results with the instrumental variables specification. The results using discrete outcomes were run using probit models. However, the results are robust to using linear probability models instead.

Looking at each outcome across the three specifications, we see that students who start at a community college had a significantly smaller likelihood of degree receipt, fewer total credits earned, and an increased likelihood of stopping out without a degree. However, as expected, upon controlling for selection, the size of the estimated penalty falls. In terms of nine-year graduation rates, the negative effect of starting at a community college falls from 0.217 (specification 1) to 0.211 when using propensity score methods (specification 2), a very minor difference. However, using the IV approach to control for unobserved differences between community college and nonselective four-year entrants consistently results in the smallest estimate of the penalty (0.145 in terms of degree completion within nine years; specification 11). For credit hours completed, over six years, there is a difference of 15.4 credits in the simple comparison (specification 4), but this drops to 14.4 credits in the PS results, and 5.6 credits in the IV estimate (specification 14). Therefore, while propensity scores likely address some small part of the selection bias in the straightforward results, the IV strategy, by dealing with unobservable differences, seems to do a better job addressing concerns about the comparability of students across types of schools. For the IV estimates, we also report overidentification tests under the null hypothesis that the instruments are valid. The test is $\chi 2$ (chi-square) distributed in the number of overidentifying restrictions.[22] As is evident in the table, the test consistently fails to

[22] Intuitively, it is a test of whether, after eliminating its correlation with choice (instrumented) and the other covariates, the educational outcomes are still correlated with the distance variables.

reject the null hypothesis in all cases except for specification 15 (stop out behavior within six years), and so we have confidence in the results.

While the penalty experienced by community college students is much smaller once accounting for bias, there is still an important difference in outcomes not explained by initial selection into two-year versus four-year institutions. Even after adjusting for selection bias, our analyses still find a consistent, albeit more modest, negative effect from beginning at a community college. Focusing on the IV estimates, the penalty becomes larger once taking into account a longer period of time: the difference in rates of degree completion are not all that different within four-years (-0.019) because few students who began at nonselective four-year universities are able to do this, but the gap grows once giving students six (-0.98) and nine years (-0.145). Likewise, the difference in total credits earned within six years is equivalent to nearly two average college courses.

This conclusion is similar to that found in earlier work (Rouse, 1995; Leigh & Gill, 2003; Sandy, Gonzalez, & Hilmer, 2006; Alfonso, 2006). However, these studies are based on cohorts that graduated high school fifteen or even twenty-five years ago. Meanwhile, there has been increasing pressure on community colleges to fill the role as an alternative pathway to a baccalaureate degree due to capacity constraints, rising admissions requirements, and escalating tuition costs at many four-year institutions. With our focus on a cohort that is the most recently available to investigate nine year outcomes (having entered college in fall 1998), we provide a timely estimate of the impact of starting at a community college.

How does the Community College Penalty differ by Student Type?

Earlier work on college access and degree receipt has pointed to important differences by race and gender in the educational attainment process. As such, we ask if the penalty of community college entry is greater for females versus males, black versus white students, and lower achieving students versus higher achieving students. To answer this question, we test three sets of interactions in Tables 4, 5, and 6. It is important to note that the comparison of estimates across strategies becomes more complicated using the interactions due to the need to take into account the overall effect, interaction coefficient, and baseline dummy variable. Panel B of each table adds these coefficients together to calculate the overall effect for each group.

Table 4 summarizes the results for gender for our two most important outcomes: baccalaureate degree completion within nine years and total credit hours completed within six years. Each of the models includes an interaction between the variable of interest (starting at a community college) and a dummy variable equal to one for female students. As shown in the table, there is a statistically significant interaction between two-year entry and being female for the simple probit and PS estimates (while negative and similar in size, the IV estimates are not statistically significant). This suggests women experience a larger penalty for initially entering a community college than men. However, because women tend to do better than men overall in terms of educational outcomes (regardless of where they attend college), the overall effects are still more negative for men, as shown in panel B. Looking at the overall effects, it is still clear that correcting for selection using PS or IV models reduces the estimate of the community college penalty for both men and women.

We next test for whether the community college penalty is greater for Black versus White students. Table 5 summarizes the results by race; the sample is limited to black and white students only so the total sample size decreases to 13,836. Each of the models includes an interaction between the variable of interest (starting at a community college) and a dummy variable equal to one for being a black student. Here we note that the interaction variable is statistically significant across all the models. In other words, a differential effect is found for black versus white students with black students experiencing worse outcomes from beginning at a community college than white students. Again, the size of the estimate decreases once accounting for selection using the PS and IV strategies. However, the overidentification tests for the IV models cause us to be wary of these results. Focusing on the PS estimates instead, the results suggest that Black students had worse educational outcomes (i.e., a lower overall likelihood of graduation) and attending a community college only exacerbated the educational gap between Black and White students. Controlling for the propensity score differences (i.e., the likelihood of initially choosing to attend a community college), black students are much less likely to complete a degree or more credits than white students. This is possibly because within a propensity score block, white students have better academic preparation for college than do black students.

Finally, we look at whether there are differences by previous academic achievement level. One might hypothesize that community colleges are the best entry point for students with weaker academic achievement in that they may specialize in providing the necessary preparation for more challenging work at the four-year level. Work by Light and Strayer (2000) suggests that lower-achieving students who attend less selective schools are more likely to complete college than their counterparts at more selective institutions. On the other hand, others have argued that in fact students perform better, regardless of their demonstrated ability, in more academically challenging environments. For example, some find that attending a more selective institution, regardless of relative academic preparation, is associated with higher graduation rates (Kane, 1998; Alon & Tienda, 2005). These hypotheses would lead to different conclusions about whether the penalty of starting at a community college differs (and in what direction) for students of different demonstrated abilities.

Table 6 summarizes the results for models where we include interactions between type of college entrance and a categorical measure of ACT score — low (0-17) and high (23-36) with the left out categories being average scores (18-22). Without controlling for bias, it appears that higher-ability students at community colleges are at a real disadvantage over average ability students. In the propensity score analysis, students with higher ACT scores are less likely to complete a degree or more credits than students with more average ACT scores. This is because within a propensity score block, students with more average ACT scores are likely to otherwise have better performance on other indicators (years of high school math or English, high school rank, etc.), or may otherwise be different on other demographics (race, parental income, etc.) and therefore, have better outcomes overall. But it may be the case that there are unobservable differences we are not catching – it may be the case that students with higher scores who enter the community colleges are less motivated, actually lack the right preparation to succeed, or have other demands on their time. Once we account for this, higher-ability students do not suffer any special penalty in terms of graduation and the interaction results are not statistically significant in terms of credits completed. On the other side, students with lower ACT scores also seem to do worse at community colleges than the average student in terms of degree completion (the results for credits completed are not statistically significant). These results are robust across the estimation strategies and lend credence to students doing better if at a nonselective four-year institution.

V. CONCLUSION

Although community colleges serve an important function in postsecondary education, comparatively little education research assesses their effectiveness or examines their students. While several studies have examined the consequences of attending a community college, many questions remain about both the goals and intentions of community college students and their post-collegiate outcomes. The increased flexibility and nontraditional patterns of attendance afforded by community colleges may weaken a student's progression through postsecondary education and make transferring to a four-year college less likely (Grubb, 1991). Furthermore, the administrative costs of transferring to a four-year college may be exceedingly prohibitive for many students.

Using a detailed dataset of individuals in a large, public higher education system, this paper provides a timely examination of the experiences of community colleges students who intend to get a four-year degree. Our paper, beyond utilizing straightforward regression models, uses two additional research strategies to account for the fact that students do not sort into community colleges versus four-year institutions randomly. Accounting for selection using both propensity score methods and instrumental variables gives us the opportunity to measure and compare differences in postsecondary outcomes in two ways. In addition, we use a longer time span of data (nine years) than any other study, have important information on transfer behavior and student degree intent, and shed light on a context that has not previously been studied.

Although the paper does not specifically test a democratization or diversion theory of community colleges, it does however ask whether, for those who divert (or perhaps detour) to the bachelor's degree via the community college, is there a cost? We find that in fact there is a cost in terms of degree completion, credit accumulation, and risk of dropping out to initially entering postsecondary study through the community college. In other words, we find a persistent community college penalty. Moreover, this penalty persists even after controlling for key student demographic and academic achievement variables. It is also evident after accounting for students' self-selection into community colleges using two different empirical methodologies. Selection into type of postsecondary institution is clearly a factor that biases straightforward estimates of the penalty of starting at a community college. We also find that the penalty of starting at two-year rather than nonselective four-year institutions varies by type of student with female students experiencing a different penalty than male students.

As noted above, these results may be a conservative estimate of the "community college penalty" due to our sample restriction that students must have taken the ACT; the results of this paper do not reflect the penalty experienced by students who want a baccalaureate degree but failed to prepare for admission into a four-year university, a group who is likely to have less academic preparation. As shown by the summary statistics in Table 1, the results are representative for more traditional-age, primarily white college students on the margin between attending a two- and four-year institution who have slightly below average ACT scores (somewhere in the 25th to 50th percentile in scores).[23]

These findings are particularly relevant in today's debates about how to improve educational attainment given the growing importance of community colleges as a pathway to a baccalaureate degree. Amidst rising demands for skilled workers coupled with the increasing numbers of individuals seeking higher education, community colleges have been forced to accommodate much of the expansion in postsecondary schooling. However, one must acknowledge that, on average, the outcomes of students who initially enter higher education through the two-year system appear to lag behind those who enter via a four-year college. Our conservative estimates suggest that these students are 14.5 percent less likely to complete a baccalaureate degree within nine years. This has significant consequences, especially for low-income and minority students who disproportionally rely on the community colleges as the primary portal for postsecondary entry. Due to the "penalty" experienced by community colleges students, caution should be exercised when designing policies that might shift enrollment patterns more towards the two-year colleges. On the other hand, because community colleges are less expensive, it is worth comparing the size of the penalty to the difference in costs at two-year versus four-year institutions. In addition, greater focus is warranted on institutional policies and programs that support community college students and help them transfer to four-year institutions to reach their intended goals of obtaining a baccalaureate degree.

---

[23] The average composite ACT score in 1997-98 was 21. Source: ACT (n.d.) *The 1998 ACT High School Profile Report - National Data.* Accessed on May 14, 2008 from http://www.act.org/news/data/98/tsum.html.

REFERENCES

Agodini, R. & Dynarski, M. (2004). Are Experiments the Only Option? A Look at Dropout Prevention Programs. *The Review of Economics and Statistics*, *86*(1), 180-194.

Adelman, C. (2005). *The Toolbox Revisited*. Washington, D.C.: U.S. Department of Education.

Alfonso, M. (2006). The impact of community college attendance on Baccalaureate Attainment. *Research in Higher Education*, *47*(8), 873-903.

Alfonso, M., Bailey, T., & Scott, M. (2005) The Educational Outcomes of Occupational Sub-baccalaureate Students: Evidence from the 1990s. *Economics of Education Review, 24*(2),197-212.

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

Anderson, G., Sun, J. C., & Alfonso, M. (2006) Effectiveness of Statewide Articulation Agreements on the Probability of Transfer: A Preliminary Policy Analysis. *Review of Higher Education*, *29*(3), 261-291.

Bailey, T. R., & Weininger, E. (2002) Performance, Graduation, and Transfer of Immigrants and Natives in the City University of New York Community Colleges. *Educational Evaluation and Policy Analysis*, *24*(4), 359-377.

Becker, G. (1993). *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education.* Chicago: The University of Chicago Press.

Brint, S., & Karabel, J. (1989). *The Diverted Dream: Community Colleges and the Promise of Educational Opportunity in America, 1900-1985*. New York: Oxford University Press.

Calcagno, J. C, Crosta, P., Bailey, T. R., & Jenkins, D. (2007) Stepping Stones to a Degree: The Impact of Enrollment Pathways and Milestones on Community College Student Outcomes. *Research in Higher Education*, *48*(7), 755-801.

Card, D. (1995) Using Geographic Variation in College Proximity to Estimate the Return to Schooling". In L.N.Christofides, E.K. Grant, & R. Swidinsky (Eds.), *Aspects of Labor Market Behaviour: Essays in Honour of John Vanderkamp*. Toronto: University of Toronto Press.

Conniffe, D., Gash, V., & O'Connell, P. (2000). Evaluating State Programmes: "Natural Experiments" and Propensity Scores. *The Economic and Social Review*, *31*(4), 283-308.

Dehejia, R. H., & Wahba. (2002). Propensity Score-Matching Methods for Nonexperimental Causal Studies. *The Review of Economics and Statistics*, *84* (1), 151-161.

Dehejia, R. H., & Wahba. (1999). Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs. *Journal of the American Statistical Association*, *94* (448), 1053-1062.

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

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

Dougherty, K., & Kienzl. G. S. (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.

Ehrenberg, R. & Smith, C. (2004) Analyzing the Success of Student Transitions from 2- to 4-year Institutions within a State. *Economics of Education Review*, *23* (1), 11-28.

Evelyn, J. (2002) "Budget Cuts Force Community Colleges to Consider Turning Away Students." *Chronicle of Higher Education, 48*(46),A25-A26, July 2002.

Grubb, W. N. (1995). Postsecondary education and the sub-baccalaureate labor market: Corrections and extensions. *Economics of Education Review*, *14*(3), 285-299.

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

Hilmer, M. J. (1997). Does community college attendance provide a strategic path to a higher quality education? *Economics of Education Review*, *16*(1), 59-68.

Hilmer, M. J. (1998). Post-secondary fees and the decision to attend a university or a community college. *Journal of Public Economics*, *67*(3), 329-348.

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

Kane, T. (1998). "Misconceptions in the Debate over Affirmative Action in College Admissions" in G. Orfield & E. Miller (Eds.), *Chilling Admissions: The Affirmative Action Crisis and the Search for Alternatives* (pp.17-31). Cambridge, MA: Harvard Education Publishing Group.

Kane, T., & Rouse, C. (1995). "Labor Market Returns to Two- and Four-Year College." *American Economic Review*, *85*(3), 600-614.

Kyung, W. (1996). In-migration of College Students to the State of New York. *The Journal of Higher Education*, *67*(3), 349-358.

Lee, V., & Frank, K. (1990). Students' Characteristics that facilitate the Transfer from Two-year to Four-year Colleges. *Sociology of Education*, *63*(3), 178-193.

Leigh, D. E., & Gill, A. M. (2003). Do community colleges really divert students from earning bachelor's degrees? *Economics of Education Review*, *22*(1), 23-30.

Leigh, D. E., & Gill, A. M. (2007). *Do Community Colleges Respond to Local Needs: Evidence from California*. Kalamazoo: W.E. Upjohn Institute for Employment Research.

Light, A., & Strayer, W. (2000). Determinants of College Completion: School Quality or Student Ability. *Journal of Human Resources*, *35*(2), 299-332.

Long, B. T. (2004). How Have College Decisions Changed Overtime? An Application of the Conditional Logistic Choice Model. *Journal of Econometrics*, *121*(1-2), 271-296.

Melguizo, T., Kienzl, G., & Alfonso, M. (2007). "Evaluating the effect of community college attendance on educational attainment: Application of propensity score matching methods." Paper presented at the American Educational Research Association Annual Meeting ,Chicago, IL.

Mills, Kay. (2006). "Enrollment Squeeze: Virginia's community colleges cope with increasing demand and a changing world." *National CrossTalk,* 14(1), Winter 2006.

Moore, C., & Shulock, N. (2007). *Beyond the Open Door: Increasing Student Success in the California Community Colleges.* Sacramento: Institute for Higher Education Leadership & Policy, California State University- Sacramento.

Morgan, S.L. (2001). Counterfactuals, Causal Effect Heterogeneity, and the Catholic School Effect on Learning. *Sociology of Education*, *74*, 341-374.

Mortenson, T. (2002). “Chance for College by Age 19 by State in 2000.” Oskaloosa, IA: Postsecondary Education OPPORTUNITY, no. 123, September.

NCES (2002). National Education Longitudinal Study of 1988: Base-Year to Fourth Follow-up Data File User's Manual (NCES 2002-323). Table C-12: NELS:88 fourth follow-up study completion rates, by selected sample strata: 2000. Accessed on April 21, 2008, from: http://nces.ed.gov/pubs2002/2002323.pdf.

Ohio Board of Regents (2001). Ohio Colleges and Universities 2001: Profile of Student Outcomes, Experiences and Campus Measures. Columbus, O.H.: Ohio Board of Regents.

Ohio Board of Regents (2002). Making the Transition from High School to College in Ohio 2002. Columbus, O.H.: Ohio Board of Regents.

Roksa, J. (2006). Does Vocational Focus of Community Colleges Hinder Students’ Educational Attainment? *The Review of Higher Education, 29*, 499-526.

Rouse, C. E. (1995). Democratization or Diversion? The Effect of Community Colleges on Educational Attainment. *Journal of Business Economics and Statistics*, *13*(2), 217-224.

Rouse, C. (1998). Do two-year colleges increase overall educational attainment? Evidence from the states. *Journal of Policy Analysis and Management*, *17*(4), 595-620.

Rosenbaum, P.R , & Rubin, D.B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, *70*, 41-55.

Rubin, D.B. (1997). Estimating Causal Effects from Large Data Sets Using Propensity Scores. *Annals of Internal Medicine*, 127(8), 757-763.

Rosenbaum, J. (2001). *Beyond College For All: Career Paths for the Forgotten Half*. New York, NY: Russell Sage Foundation.

Sandy, J., Gonzalez, A., &. Hilmer, M. (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.

Smith, J.A., & Todd, P. E. (2004). Does matching overcome LaLonde’s critique of nonexperimental estimators? *Journal of Econometrics*, *125*, 305-353

Stephan, J. L. and Rosenbaum, J. (2006). "Do Some Colleges Improve Students’ Chances of Completing Degrees? How Propensity Scores Change the Question" Paper presented at the annual meeting of the American Sociological Association, Montreal Convention Center, Montreal, Quebec, Canada (August 2006).

Surette, B. J. (2001). Transfer from two-year to four-year college: An analysis of gender differences. *Economics of Education Review*, *20*(2), 151–163.

Whitaker, D., & Pascarella, E. (1994). Two-Year College Attendance and Socioeconomic Attainment: Some Additional Evidence. *Journal of Higher Education*, *65*(2), 194-210.

Table 1: Descriptive Statistics

Sample: Fall 1998 entering cohort age 17-20 who had intent to complete a baccalaureate degree

	Public, Four-year Selective Universities	Public, Four-year Nonselective Universities	Public, Two-year Community Colleges
Age in 1998	18.351 (0.4995)	18.432 (0.5898)	18.425 (0.6446)
Percent Female	0.5500	0.5387	0.5498
Percent White	0.8655	0.7991	0.8973
Percent African American	0.0730	0.1178	0.0544
Percent Hispanic	0.0151	0.0171	0.0142
Percent Asian	0.0245	0.0148	0.0073
High School GPA	3.297 (0.4467)	2.959 (0.5901)	2.816 (0.5871)
Grades in HS Math	3.204 (0.7829)	2.758 (0.9905)	2.548 (1.0249)
Grades in HS English	3.326 (0.7195)	2.959 (0.8589)	2.805 (0.9084)
ACT Composite Score (36 maximum)	23.553 (3.838)	20.730 (4.2603)	19.550 (3.7237)
ACT Math	23.515 (4.480)	20.488 (4.6840)	19.198 (3.8956)
ACT English	22.938 (4.397)	19.965 (4.8705)	18.692 (4.5223)
In Math Remediation	0.0807 [11,255]	0.2749 [7421]	0.4639 [6889]
In English Remediation	0.0237 [11,255]	0.2372 [7421]	0.2700 [6889]
Percent Full-time	0.8188	0.6708	0.5550
Stopped Out within 1 year	0.0431	0.1103	0.1642
Stopped Out within 2 years	0.0852	0.1886	0.2761
Stopped Out within 6 years	0.1810	0.3406	0.4392
Total Credit Hours after 6 yrs	110.893 (43.058)	85.218 (52.016)	66.230 (48.247)
Rcvd Bach. Deg. within 4 yrs	0.3647	0.1525	0.0564
Rcvd Bach. Deg. within 6 yrs	0.6900	0.4372	0.2118
Rcvd Bach. Deg. within 9 yrs	0.7291	0.4972	0.2587
Observations	13,683	7,422	7,388

Source: Ohio Board of Regents (OBR) Higher Education Information (HEI) System: Student Entrance, Student Enrollment, and Course Enrollment files along with test score and student background information from ACT, Inc.

Notes: The sample consists of the fall 1998 entering cohort age 17-20 who had intent to complete a baccalaureate degree. Due to the estimation techniques used, we also limit the sample to students who took the ACT (because of the information it provides on background) and Ohio residents (because we use residential proximity to an Ohio college as an instrumental variable). Standard deviations are shown in the parentheses. If the entire sample was not used in the calculation of the statistic, the observations are shown in brackets. In Ohio, the selective, public institutions require a certain academic standard but are not considered highly selective by national norms; in contrast, the nonselective, four-year universities are open admissions schools.

Table 2: The Role of the Initial Institution on Degree Completion
Probit Models with Marginal Effects reported
Sample: Fall 1998 entering cohort age 17-20 who had intent to complete a baccalaureate degree

	ALL FOUR- AND TWO-YEAR INSTITUTIONS				
	Outcome: Completed a Bachelor's Degree within 6 Years				*Outcome:* Completed a Bachelor's Degree within 9 Years
	Baseline	Adding Demographics	Adding Parents' Income	Adding ACT percentile	
	(1)	(2)	(3)	(4)	(5)
Initially Attend 2yr	-0.4259** (0.0069)	-0.4274** (0.0071)	-0.4175** (0.0072)	-0.3626** (0.0081)	-0.3544** (0.0084)
Initially Attend Nonselective 4yr	-0.2136** (0.0076)	-0.1920** (0.0078)	-0.1799** (0.0079)	-0.1415** (0.0082)	-0.1266** (0.0083)
Pseudo R^2	0.09	0.11	0.12	0.14	0.14
Observations	28,493	28,493	28,493	28,493	28,493
	NONSELECTIVE FOUR- AND TWO-YEAR INSTITUTIONS ONLY				
	Outcome: Completed a Bachelor's Degree within 6 Years				*Outcome:* Completed a Bachelor's Degree within 9 Years
	Baseline	Adding Demographics	Adding Parents' Income	Adding ACT percentile	
	(6)	(7)	(8)	(9)	(10)
Initially Attend 2yr	-0.2254** (0.0075)	-0.2409** (0.0075)	-0.2396** (0.0076)	-0.2023** (0.0078)	-0.2174** (0.0081)
Pseudo R^2	0.05	0.07	0.07	0.12	0.11
Observations	14,810	14,810	14,810	14,810	14,810

* significant at 10% ** significant at 5%

Notes: Standard errors are shown in parentheses. Due to the estimation techniques used, the sample is also limited to ACT-takers and Ohio residents. The baseline only includes the variables for where the student initially attended (the shown coefficients). The second column adds the following demographic controls: a dummy variable for female students (baseline: male students), dummy variables for Black, Hispanic, Asian, and Native American students (baseline: White students), age, and age-squared. The third column adds parental income, parental income-squared, and a dummy indicating missing values for parental income. The fourth column adds the math and English ACT scores. The last column includes all the controls (the same as the fourth column).

Table 3: The Role of the Initial Institution – Comparing Estimation Strategies
Probit Models with Marginal Effects reported

Sample: Students age 17-20 who entered a two- or nonselective four-year college during fall 1998 and had intent to complete a baccalaureate degree

Outcome	Completed Bachelor's Degree 9 yrs	Completed Bachelor's Degree 6 yrs	Completed Bachelor's Degree 4 yrs	Total Credit Hrs 6 years	Stop out within 6 years
PANEL A: PROBIT RESULTS					
	(1)	(2)	(3)	(4)	(5)
Initially Attend 2yr	-0.2174**	-0.2023**	-0.0693**	-15.3948**	0.0796**
	(0.0081)	(0.0078)	(0.0046)	(0.8085)	(0.0083)
Pseudo R^2/ R^2	0.11	0.12	0.13	0.1285	0.04
PANEL B: PROPENSITY SCORE PROBIT RESULTS					
	(6)	(7)	(8)	(9)	(10)
Initially Attend 2yr	-0.2110**	-0.1953**	-0.0642**	-14.4028**	0.0734**
	(0.0082)	(0.0079)	(0.0045)	(0.8035)	(0.0083)
Pseudo R^2/ R^2	0.12	0.13	0.14	0.1465	0.04
PANEL C: INSTRUMENTAL VARIABLES PROBIT RESULTS					
	(11)	(12)	(13)	(14)	(15)
Initially Attend 2yr	-0.1451**	-0.0983**	-0.0194**	-5.5882**	0.0488**
	(0.0170)	(0.0164)	(0.0090)	(1.6446)	(0.0169)
Overidentification test ($\chi^2_{d.f.=1}$)	0.371	1.227	1.760	0.204	7.801
Overid. test P-value	0.5427	0.2681	0.1846	0.6517	0.0052
Pseudo R^2/ R^2	0.08	0.08	0.11	0.1198	0.03

* significant at 10% ** significant at 5%

Notes: Standard errors are shown in parentheses. Due to the estimation techniques used, the sample is also limited to ACT-takers and Ohio residents. All regressions have 14,810 observations. The results for "Total Credit Hours" were estimated using OLS instead of probit models. R-squares are displayed for the "Total Credit Hours" results. All others are Pseudo R-squares. The number of blocks with the propensity score estimates is 10. Each regression includes the following controls: a dummy variable for female students (baseline: male students), dummy variables for Black, Hispanic, Asian, and Native American students (baseline: White students), age, age-squared, parental income, parental income-squared, a dummy indicating missing values for parental income, and the student's math and English ACT scores. The results using discrete outcomes were run using probit models. However, the results are robust to using linear probability models instead.

Table 4: The Role of Initial Institution on Educational Outcomes by Gender – IV Results
Probit Models with Marginal Effects reported

Sample: Students age 17-20 who entered a two- or nonselective four-year college during fall 1998 and had intent to complete a baccalaureate degree

	Completed Bachelor's Degree within 9 yrs			Total Credit Hrs (6 years)		
	Probit Results (1)	PS Results (2)	IV Results (3)	Probit Results (4)	PS Results (5)	IV Results (6)
A. ESTIMATED COEFFICIENTS						
Initially Attend 2yr	-0.1915** (0.0121)	-0.1841** (0.0122)	-0.1165** (0.0256)	-12.8725** (1.1809)	-11.8093** (1.1724)	-2.9279 (2.4406)
Initially Attend 2yr * Female Dummy Variable	-0.0483** (0.0163)	-0.0499** (0.0164)	-0.0503 (0.0326)	-4.6203** (1.5770)	-4.7436** (1.5619)	-4.7042 (3.1661)
Female Dummy Variable	0.0889** (0.0117)	0.0908** (0.0117)	0.0904** (0.0187)	11.3608** (1.1533)	11.5337** (1.1425)	11.4751** (1.8036)
Overidentification test ($\chi^2_{d.f.=3}$)	---	---	3.713	---	---	0.139
Overid. test P-value	---	---	0.2942			0.9867
Pseudo R^2/ R^2	0.11	0.12	0.08	0.1290	0.1471	0.1201
B. OVERALL EFFECT OF STARTING AT A COMMUNITY COLLEGE						
Male Students	-0.1915	-0.1841	-0.1165	-12.8725	-11.8093	-2.9279
Female Students	-0.1509	-0.1432	-0.0764	-6.1320	-5.0192	3.8430

* significant at 10%   ** significant at 5%

Notes: Standard errors are shown in parentheses. The sample is students age 17-20 who entered a nonselective two- or four-year college during fall 1998 and had intent to complete a baccalaureate degree. Due to the estimation techniques used, the sample is also limited to ACT-takers and Ohio residents. All regressions have 14,810 observations. The results for "Total Credit Hours" were estimated using OLS instead of probit models. R-squares are displayed for the "Total Credit Hours" results. All others are Pseudo R-squares. The number of blocks with the propensity score estimates is 10. Each regression includes the following controls: a dummy variable for female students (baseline: male students), dummy variables for Black, Hispanic, Asian, and Native American students (baseline: White students), age, age-squared, parental income, parental income-squared, a dummy indicating missing values for parental income, and the student's math and English ACT scores.

Table 5: The Role of Initial Institution on Educational Outcomes by Race (Black and White Students Only) – Probit Models with Marginal Effects reported

Sample: Students age 17-20 who entered a two- or nonselective four-year college during fall 1998 and had intent to complete a baccalaureate degree

	Completed Bachelor's Degree within 9 yrs			Total Credit Hrs (6 years)		
	Probit Results (1)	PS Results (2)	IV Results (3)	Probit Results (4)	PS Results (5)	IV Results (6)
A. ESTIMATED COEFFICIENTS						
Initially Attend 2yr	-0.2112** (0.0087)	-0.2043** (0.0088)	-0.1281** (0.0182)	-14.7258** (0.8720)	-13.6650** (0.8665)	-3.6118** (1.7655)
Initially Attend 2yr * Black Dummy Variable	-0.1203** (0.0343)	-0.1162** (0.0344)	-0.1590** (0.0657)	-8.5411** (2.9982)	-8.5039** (2.9950)	-18.0057** (6.7629)
Black Student Dummy Variable	-0.0547** (0.0180)	-0.2914** (0.0178)	-0.0133 (0.0313)	-8.5703** (1.8182)	-45.4329** (3.0703)	-2.1279 (2.7534)
Overidentification test ($\chi^2_{d.f.=3}$)	---	---	10.359	---	---	8.272
Overid. test P-value	---	---	0.0158	---	---	0.0407
Pseudo R^2/ R^2	0.11	0.12	0.08	0.1306	0.1490	0.1203
B. OVERALL EFFECT OF STARTING AT A COMMUNITY COLLEGE						
White Students	-0.2112	-0.2043	-0.1281	-14.7258	-13.6650	-3.6118
Black Students	-0.3862	-0.6119	-0.3004	-31.8372	-67.6018	-23.7454

* significant at 10% ** significant at 5%

Notes: Standard errors are shown in parentheses. The sample is students age 17-20 who entered a nonselective two- and four-year college during fall 1998 and had intent to complete a baccalaureate degree. Due to the estimation techniques used, the sample is also limited to ACT-takers and Ohio residents. All regressions have 13,836 observations due to the fact non-white and non-black students were dropped from the sample for this analysis. The results for "Total Credit Hours" were estimated using OLS instead of probit models. R-squares are displayed for the "Total Credit Hours" results. All others are Pseudo R-squares. The number of blocks with the propensity score estimates is 10. Each regression includes the following controls: a dummy variable for female students (baseline: male students), dummy variables for Black, Hispanic, Asian, and Native American students (baseline: White students), age, age-squared, parental income, parental income-squared, a dummy indicating missing values for parental income, and the student's math and English ACT scores.

Table 6: The Role of Initial Institution on Outcomes by Preparation – IV Results
Probit Models with Marginal Effects reported

Sample: Students age 17-20 who entered a two- or nonselective four-year college during fall 1998 and had intent to complete a baccalaureate degree

	Completed Bachelor's Degree within 9 yrs			Total Credit Hrs (6 years)		
	Probit Results (1)	PS Results (2)	IV Results (3)	Probit Results (4)	PS Results (5)	IV Results (6)
A. ESTIMATED COEFFICIENTS						
Initially Attend 2yr	-0.1981** (0.0116)	-0.1871** (0.0117)	-0.1308** (0.0231)	-13.6647** (1.1816)	-12.0216** (1.1596)	-3.5609 (2.3114)
Initially Attend 2yr * High ACT Score	-0.0794** (0.0186)	-0.0621** (0.0192)	-0.0121 (0.0403)	-11.4030** (1.9837)	-9.0981** (1.9530)	-6.1703 (4.0394)
Initially Attend 2yr * Low ACT Score	-0.0414** (0.0198)	-0.0423** (0.0201)	-0.0745* (0.0387)	-1.6363 (1.9221)	-1.3450 (1.8896)	-4.7433 (3.8159)
High ACT Score Dummy Variable	0.1663** (0.0135)	-0.0361** (0.0163)	0.1488** (0.0210)	17.0388** (1.3107)	-4.9744** (1.6080)	16.4880** (2.0191)
Low ACT Score Dummy Variable	-0.0962** (0.0139)	0.0445** (0.0173)	-0.0826** (0.0241)	-11.7951** (1.4546)	4.6348** (1.6327)	-10.8974** (2.3609)
Overidentification test ($\chi^2_{d.f.=5}$)	---	---	5.800	---	---	3.742
Overid. test P-value	---	---	0.3262	---	---	0.5871
Pseudo R^2/ R^2	0.09	0.12	0.05	0.1068	0.1460	0.0957
B. OVERALL EFFECT OF STARTING AT A COMMUNITY COLLEGE						
High ACT	-0.1112	-0.2853	0.0059	-8.0289	-26.0941	6.7568
Middle ACT	-0.1981	-0.1871	-0.1308	-13.6647	-12.0216	-3.5609
Low ACT	-0.3357	-0.1849	-0.2879	-27.0961	-8.7318	-19.2016

* significant at 10% ** significant at 5%

Notes: Standard errors are shown in parentheses. The sample is students age 17-20 who entered a nonselective two- and four-year college during fall 1998 and had intent to complete a baccalaureate degree. Due to the estimation techniques used, the sample is also limited to ACT-takers and Ohio residents. All regressions have 14,810 observations. The results for "Total Credit Hours" were estimated using OLS instead of probit models. R-squares are displayed for the "Total Credit Hours" results. All others are Pseudo R-squares. The number of blocks with the propensity score estimates is 10. "High ACT" is a dummy variable for having a composite ACT score of 23 to 36; "Middle ACT" is for scores 18 to 22; and "Low ACT is for scores 17 and below. Each regression includes the following controls: a dummy variable for female students (baseline: male students), dummy variables for Black, Hispanic, Asian, and Native American students (baseline: White students), age, age-squared, parental income, parental income-squared, a dummy indicating missing values for parental income, and dummy variables for having a high ACT score (23-36) or low ACT score (17 or below).

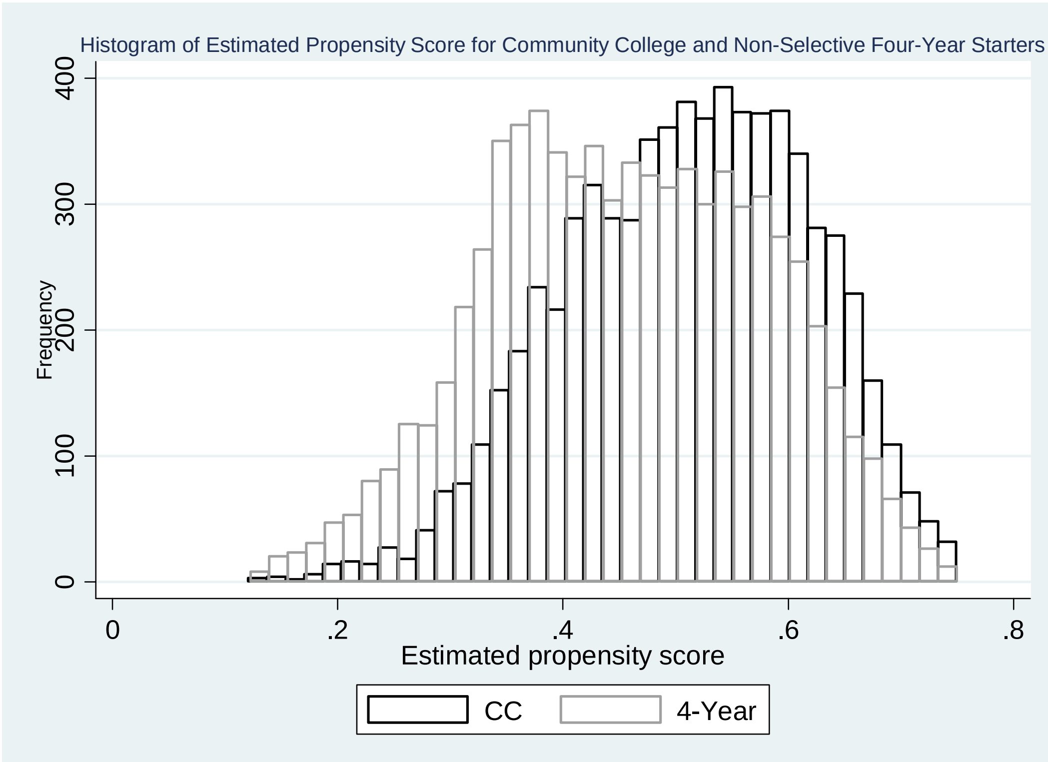
Appendix Table 1: Logit Coefficients from the Model Estimating the Probability of Entering a Public Two-Year College

	Community College versus Nonselective Four-Year University
Female	0.0234 (0.0363)
Black	-1.2707*** (0.0682)
Hispanic	-0.3737** (0.1379)
Asian	-0.8822*** (0.1759)
Native Am	-0.0360 (0.3414)
Age	-0.1256*** (0.0283)
Parent Income	-2.27e-06** (7.82e-07)
Parent Income Missing Dummy	0.0086 (0.0709)
High School GPA	0.3533 (0.2430)
High School GPA-squared	-0.1244** (0.0438)
High School GPA Rank	0.0833 (0.1090)
Years of High School Math	-0.0961** (.0324)
Years of High School English	-0.0911~ (0.0513)
High School Math Grades	1.5928*** (0.1281)
Missing HS Math grades dummy	0.4193*** (0.1226)
High School English Grades	0.0039 (0.0340)
Missing HS English grades dummy	0.3444** (0.1702)
ACT percentile	-0.0112*** (0.0008)
Private High School	-0.3452*** (0.0539)
Constant	3.3205 (0.6629)
R-squared	.0672
Observations	14,819

~ significant at 10% * significant at 5% ***significant at .01%

Notes: The sample is limited to fall 1998 first-time students age 17-20 who had four-year degree intent, took the ACT, and were Ohio residents. Standard errors are shown in the parentheses.

Appendix Figure 1: Histogram of the Estimated Propensity Score for the Two-year versus Nonselective Four-year Starters



Appendix Table 2: Mean Propensity Scores for Two-year Entrants (treated) and Nonselective Four-year Entrants (control) by Strata (Blocks)

Blocks	Mean Propensity Score		Number of Observations		T-Test Mean(C)-Mean (T)
	Community College (Treatment)	Four-Year (Control)	Community College (Treatment)	Four-Year (Control)	
Block 1	.2099	.2076	79	315	-0.5738
Block 2	.2672	.2654	50	231	-1.2420
Block 3	.3286	.3280	501	1190	-0.5486
Block 4	.3679	.3673	197	313	-1.4686
Block 5	.3905	.3898	422	674	-1.3004
Block 6	.4215	.4221	580	637	1.1514
Block 7	.4704	.4692	1227	1210	-1.6696
Block 8	.5501	.5488	2274	1845	-1.4120
Block 9	.6191	.6184	745	542	-1.0513
Block 10	.6748	.6746	812	454	-0.1740

Appendix Table 3: First stage IV Results – Probit Models with Marginal Effects reported
Sample: Students age 17-20 who entered a two- or nonselective four-year college during fall 1998 and had intent to complete a baccalaureate degree

	Choose Two-year College
Distance to closest Two-year College (per 100 miles)	-0.0942** (0.0458)
Distance to closest Nonselective Four-year University (per 100 miles)	1.3510** (0.0236)
Female	-0.0134 (0.0098)
Black	-0.1698** (0.0168)
Hispanic	-0.0201 (0.0357)
Asian	-0.1747** (0.0420)
Native American	-0.0289 (0.0970)
Age	-3.8844** (0.3523)
Age-squared	0.1038** (0.0095)
Parent Income (per $1,000)	0.0006 (0.0008)
Parent Income-squared (per $1,000,000)	-0.0000 (0.0000)
Parent Income Missing Dummy	0.0398 (0.0246)
ACT Math Score	-0.0188** (0.0015)
ACT English Score	-0.0096** (0.0013)
Test of joint significance of IVs	3263.11
Pseudo R-squared	0.24
Observations	14,810

* significant at 10%   ** significant at 5%
Notes: Standard errors in parentheses. The variables listed above are the instruments for the selection of a two-year college (over a nonselective four-year university). The regression also includes the following exogenous variables: a dummy variable for female students (baseline: male students), dummy variables for Black, Hispanic, Asian, and Native American students (baseline: White students), age, age-squared, parental income, parental income-squared, a dummy indicating missing values for parental income, and the student's math and English ACT scores.