THE PROMISES AND PITFALLS OF MEASURING COMMUNITY COLLEGE QUALITY

Michal Kurlaender and Scott Carrell
University of California, Davis

Jacob Jackson
Public Policy Institute of California

February 2015

Please Do Not Cite without Authors' Permission

This research was supported in part by grants from the Bill and Melinda Gates Foundation. We thank the California Community College Chancellor's Office and the California Department of Education for their assistance with data access. Opinions reflect those of the authors and do not necessarily reflect those of the granting agencies or the state agencies providing data.

Identifying college quality has been a key element of the Obama Administration's efforts to increase accountability in higher education. In 2013, the White House launched the College Scorecard with the goal of providing students and their families information about the "cost, value, and quality" of specific colleges in order to make more informed decisions.[1] Beyond transparency, the Administration is also pushing for performance-based funding in higher education. [2] Specifically, President Obama's proposal aims (by 2018) to tie federal aid to a rating system of colleges based on affordability, student completion rates, and the earnings of graduates.

There has been much discussion of these ratings, with skepticism about the quality of the data utilized for the ratings and whether, as the President of the University of California system Janet Napolitano states, "criteria can be developed that are in the end meaningful."[3] Admittedly, policymakers have recognized that there are a host of issues in developing the accountability metrics, and as such have solicited feedback on the college ratings methodology.

Among the many critiques of the rating systems is whether or not it is reasonable to compare institutions that are quite different from one another in terms of the institutional goals and the student populations served. Some have noted that even if scorecard rankings are adjusted for institutional and/or individual differences across campuses, there will still be biases that favor elite institutions and institutions that serve

---

[1] http://www.whitehouse.gov/issues/education/higher-education/college-score-card

[2] http://www.whitehouse.gov/the-press-office/2013/08/22/fact-sheet-president-s-plan-make-college-more affordable-better-bargain-

[3] Napolitano, University of California president, 'deeply skeptical' of Obama college rating plan, December 6, 2013, Nick Anderson, The Washington Post. Available at: http://www.washingtonpost.com/local/education/napolitano-uc-president-deeply-skeptical-of-key-assumption-in-obama-college-rating-plan/2013/12/06/f4f505fa-5eb8-11e3-bc56-c6ca94801fac_story.html

more traditional college students (Gross 2013).[4] Related, others worry that a rating system, particularly one that is tied to performance is "antithetical" to the open access mission of community colleges (Fein 2013).[5]

The idea of performance-based accountability may be novel in higher education, but in K-12 it has been at the heart of both federal and state accountability systems, which developed—albeit to varying success—structures to grade K-12 schools on a variety of performance measures. Long before state and federal accountability systems took hold, school leaders and the research community were preoccupied with understanding the unique effects of schools on individual outcomes. Nearly fifty years after the Coleman Report there have been many scholarly efforts to isolate the specific contribution of schools on student outcomes, controlling for individual and family characteristics.

Several studies since the canonical Coleman Report, which concluded that the differences between K-12 schools account for only a small fraction of differences in pupil achievement, have found that school characteristics explain less than 20 percent of the variation in student outcomes, though one study concludes that as much as 40 percent of the differences in student achievement is attributable to schools, even after taking into account students' family background (Startz 2012; Borman and Dowling 2010; Rumberger and Palardy 2005; Rivkin, Hanushek, and Kain 2005; Goldhaber, Liddle, Theobald, and Walch 2010). In higher education, however, school effects have primarily focused on college selectivity, or have been constrained by existing aggregate data and

[4] Ratings Are Not So Easy, August 23, 2013, By Karen Gross. Inside Higher Ed, Available at: https://www.insidehighered.com/views/2013/08/23/obamas-ratings-system-may-be-difficult-pull-essay

[5] Performance Funding Goes Federal, August 23, 2013, By Paul Fain. Inside Higher Ed, Available at: https://www.insidehighered.com/news/2013/08/23/higher-education-leaders-respond-obamas-ambitous-ratings-system-plan

small samples.

In this paper we explore the community college (institutional) effect on student outcomes in the nation's largest public two-year higher education system—the California Community College system. We seek to know whether there are significant differences in student outcomes across community college campuses after adjusting for observed student differences and potential unobserved determinates that drive selection. Additionally, we ask whether college rankings based on unadjusted mean differences across campuses provide meaningful information. To do so, we leverage a unique administrative dataset that links community college students to their K-12 records in order to control for key student inputs.

Results show that there are meaningful differences in student outcomes across the 108 California Community Colleges in our sample, after adjusting for differences in student inputs. For example, our lower bound estimates show that going from the 10^{th} to 90^{th} percentile of campus quality is associated with a 3.32 (34.3 percent) increase in student transfer units earned, an 0.07 (9.6 percent) increase in the probability of persisting, an 0.09 (40.7 percent) increase in the probability of transferring to a four-year college, and an 0.08 (27.1 percent) increase in the probability of completion. We also show that college rankings based on unadjusted mean differences can be quite misleading. After adjusting for differences across campus, the average school rank changed by over thirty ranks. As such, our results suggest that policymakers wishing to rank schools based on quality should adjust such rankings for differences in student-level inputs across campuses.

In the Background section of the paper that follows we first review the literature on college quality and then provide a description of the setting in which we conduct our study. In Section III we describe our data and measures, followed by a description of our empirical strategy. In Section IV we present our results, highlighting important implications for college quality rankings. Finally in Section V we conclude.

II. BACKGROUND

Prior Literature

Prior research on college quality has focused largely on more selective four-year colleges and universities, and on the relationship between college quality and graduates' earnings. There are good reasons for students to want to attend elite private and public universities. More selective institutions appear to have a higher payoff in terms of persistence to degree completion (Alon and Tienda 2005; Bowen, Chingos, and McPherson 2009; Small and Winship 2007; Long 2008), graduate or professional school attendance (Mullen, Goyette, and Soares 2003) and earnings later in life (Black and Smith 2006; Hoekstra 2009; Long 2008; Monks 2000). However, empirical work on the effect of college quality on earnings is a bit more mixed (Brand and Halaby 2006; Dale and Krueger 2002; Hoekstra, 2009; Hoxby 2009).

The difficulty in establishing a college effect results from the non-random selection of students into colleges of varying qualities (Black and Smith, 2004). Namely, the characteristics that lead students to apply to particular colleges may be the same characteristics that lead to better post-enrollment outcomes. Prior work has addressed this challenge largely through conditioning on key observable characteristics of students, namely academic qualifications. To more fully address self-selection, Dale and Krueger

(2002; 2012) adjust for the observed set of institutions to which students submitted an application. Dale and Krueger (2002) argue that the application set reflects students' perceptions (or 'self-revelation' in Dale and Kruger's terms) about their own academic potential; students who apply to more selective colleges and universities do so because they believe they can succeed in such environments. They find relatively small differences in outcomes for students who attended elite universities compared to those who were admitted, but chose to attend a less selective university. Cunha and Miller (2014) examine institutional differences in student outcomes across Texas's 30 traditional four-year public colleges. Their results show that controlling for student background characteristics (e.g., race, gender, free lunch, SAT score, etc.), the quality of high school attended, and application behavior significantly reduces the mean differences in average earned income, persistence and graduation across four-year college campuses. However, recent papers that exploit a regression discontinuity approach in the probability of admissions, find larger positive returns to attending a more selective university (Hoekstra, 2009; Anelli, 2014).

Community colleges are the primary point of access to higher education for many Americans, yet research on quality differences between community colleges has been sparse. The multiple missions and goals of community colleges have been well documented in the academic literature (Rosenbaum 2001; Dougherty 1994; Grubb 1991, 1996; Brint and Karabel 1983). Community colleges have also captured the attention of policymakers concerned with improving workforce shortages and the overall economic health of the nation.[6] The Obama Administration identified community colleges as key drivers in the push to increase the stock of college graduates in the U.S. and to raise the

[6] See White House Summit on Community Colleges

skills of the American workforce. “It’s time to reform our community college so that they provide Americans of all ages a chance to learn the skills and knowledge necessary to compete for the jobs of the future,” remarked President Obama at a White House Summit on Community Colleges.[7]

The distinct mission and open-access nature of community colleges and the diverse goals of the students they serve make it difficult to assess differences in quality across campuses. First, it is often unclear which outcomes should actually be measured (Bailey, Calcagno, Jenkins, Leinbach, and Kienzl 2006). Moreover, selection issues into community colleges may differ from those between four-year institutions. Nevertheless, community college quality has been a key component of the national conversation about higher education accountability. This paper is not the first to explore institutional quality differences among community colleges. A recent paper explored variation in success measures across North Carolina’s 58 community colleges, and finds that conditional on student differences, colleges were largely indistinguishable from one another in degree receipt or transfer coursework, save for the differences between the very top and very bottom performing colleges (Clotfelter, Ladd, Muschkin, and Vigdor 2013). Other papers have looked at the role of different institutional inputs as proxies for institutional quality. In particular, Stange (2012) exploits differences in instructional expenditures per student across community colleges and finds no impact on student attainment (degree receipt or transfer). This finding corroborates with Calcagno and colleagues (2008); however, they identify several other institutional characteristics that do influence student outcomes. Specifically, larger enrollment, more minority students, and more part-time faculty are

[7] WhiteHouse.gov/CommunityCollege

associated with lower degree attainment and lower 4-year transfer rates (Calcagno, Bailey, Jenkins, Kienzl, and Leinbach, 2008).

In this paper we explore institutional effects of community colleges in the state with the largest public two-year community college system, utilizing a unique administrative dataset that links students' K-12 data to postsecondary schooling at community college.

Setting

California is home to the largest public higher education system, including its 112-campus community college system. Two-thirds of all California college students attend a community college. The role of community colleges as a vehicle in human capital production was the cornerstone of California's 1960 Master Plan for Higher Education, which stipulated that the California Community Colleges are to admit "any student capable of benefiting from instruction."[8] Over the years, the California community colleges have grown and have been applauded for remaining affordable, open-access institutions. However, the state's community colleges are also continually criticized for producing weak outcomes, in particular low degree receipt and transfer rates to four-year institutions (Shulock and Moore 2007, Sengupta and Jepsen 2006).

---

[8] California Master Plan for Higher Education. Available at:
http://www.ucop.edu/acadinit/mastplan/MasterPlan1960.pdf
The Master Plan articulated the distinct functions of each of the State's three public postsecondary segments. The University of California (UC) is designated to as the state's primary academic research institution and is reserved for the top one eighth of the State's graduating high school class. The California State University (CSU) is primarily to serve the top one-third of California's high school graduating class in undergraduate training, and graduate training through the master's degree, focusing primarily on professional training such as teacher education. Finally, the California Community Colleges are to provide academic and instruction for students through the first two years of undergraduate education (lower division), as well as provide vocational instruction, remedial instruction, English as a Second Language courses, adult noncredit instruction, community service courses, and workforce training services.

Several years before Obama's proposed college scorecard, California leaders initiated greater transparency and accountability in performance through the Student Success Act, signed into law by Governor Brown in 2012. Among the components of this Act is an accountability scorecard. California Community Colleges' Student Success Scorecard tracks several key dimensions in student success: remedial course progression rate; persistence rates; completion of a minimum of 30 units (roughly equivalent to one year of fulltime enrollment status); sub-baccalaureate degree receipt and transfer status, and certificate, degree or transfer among CTE students. California's Student Success Scorecard is not focused on comparing *between* institutions, rather on performance improvement overtime *within* institutions. Nevertheless, policymakers desire critical information about the effectiveness of the postsecondary system to improve human capital production in the state and to increase postsecondary degree receipt.

In 2013, the California Community College (CCC) system served over 2.5 million students from a tremendous range of demographic and academic backgrounds. California's community colleges are situated in urban, suburban and rural areas of the state, and their students come from public high schools that are both among the best and among the worst in the nation. California is an ideal state to explore institutional differences at community colleges because of the large number of institutions present, and because of the larger governance structure of the CCC system and its articulation to the State's public four-year colleges. Moreover, the diversity of California's community college population reflects the student populations of other states in the U.S. and the mainstream public two-year colleges that educate them. Given the diversity of California's students and public schools, and the increasing diversity of students entering the nation's colleges and universities,[9] we believe that other states can learn important lessons from California's public postsecondary institutions.

III. RESEARCH DESIGN

Data

To explore institutional differences between community colleges, we employ an administrative dataset that links four cohorts of California high school juniors to the California Community College system. These data were provided by the California Community College Chancellor's Office and the California Department of Education. Since California does not have an individual identifier that follows students from K-12 to postsecondary schooling we linked all transcript and completion data for four first time freshmen fall-semester cohorts (2004-2008) aged 17-19 enrolled at a California community college with the census of California 11th grade students with standardized test score data. The match, performed on name and birth date, high school attended, and cohort, initially captured 69 percent of first-time freshmen aged 17-19 enrolled at a California community college (consistent with similar studies conducted by the California Community College Chancellor's Office matched to K-12 data).[10]

The California Community Colleges are an open access system, where any student can take any number of courses at any time, including, for example, while

[9] Between 2007 and 2018 the number of students enrolled in a college or university is expected to increase by 4 percent for whites but by 38 percent for Hispanics, 29 percent for Asian/Pacific Islanders and 26 percent for African Americans (Hussar and Bailey 2009).
[10] Our match rates may be the result of several considerations. First, the name match occurred on the first three letters of a student's first name and last name, leading to many duplicates. Students may have entered different name or birthday at the community college. Students may have omitted information at either system. Second, the denominator may also be too high; not all community college students attended California high schools. Finally, students who did attend a California high school, but did not take the 11th grade standardized tests were not included in the high school data.

concurrently enrolled in high school, or the summer before college for those who intend to start as first time freshman at a four year institution. In addition, community colleges serve multiple goals, including facilitating transfers to four-year universities, sub-baccalaureate degree and certificate, career and technical education, basic skills instruction, and supporting lifelong learning. We restrict the sample for our study to first time freshman at the community college, of traditional age. We built cohorts of students who started in the summer or fall within one year of graduating high school, who attempted more than two courses (six units) in their first year, and had complete high school test and demographic information. This sample contains 254,865 students across 108 California community college campuses.[11]

Measures

We measure four outcomes intended to capture community college success in the short term through credit accumulation and persistence into year two, as well as through degree/certificate receipt and four-year transfer. First, we measure how many transferrable units a student completes during her first year. This includes units that are transferrable to California's public four-year universities (the University of California System and the California State University system) that were taken at any community college. Second, we measure whether a student persists to her second year of community college. This outcome indicates whether a student attempts any units in the fall semester after her first year at any community college in California. Third, we measure whether a student ever transfers to a four-year college. Using National Student Clearinghouse data

[11] We excluded the three campuses that use the quarter system, as well as three adult education campuses. Summer students were only allowed in the sample if they took enough units in their first year to guarantee they also took units in the fall.

that the CCC Chancellor's office linked with their own data, we are able to tell whether a student transferred to a four-year college at any point after attending a California community college. Lastly, we measure degree/certificate completion at a community college. This measure indicates whether a student earned an AA degree, or a 60 unit certificate, or transferred to a four-year university. It is important to note that these outcomes represent only a few of the community college system's many goals, and as such are not meant to be an exhaustive list of how we might examine community college quality or effectiveness.

Our data are unique in that we have the ability to connect a student's performance and outcomes at community college with their high school data. As community colleges are open access, students do not submit transcripts from their high school, and have not necessarily taken college entrance exams such as the SAT or ACT to enter. As a result, community colleges often know very little about their students' prior educational backgrounds. Researchers interested in understanding the community college population often face the same constraints. Examining the outcomes of community colleges' without considering the educational backgrounds of the students enrolling in that college may confound college effects with students' self-selection.

To address ubiquitous selection issues, we are able to adjust our estimates of quality for important background information about a student's high school academic performance. We measure a student's performance on the 11th grade English and mathematics California Standardized Tests (CSTs).[12] We are also able to determine which math course a student took in 11th grade. In addition we measure race/ethnicity, gender, and parent education levels from the high school file as sets of binary variables.

[12] We include CST scaled scores, which are approximately normally distributed across the state.

To account for high school quality, we include the Academic Performance Index (API) of high school attended. Importantly, as students are enrolling in community college, they are asked about their goals for attending community college. Students can pick from an extensive list of 15 choices, including: transfer with an associate's degree, transfer without an associate's degree, vocation certification, discover interests, improve basic skills, undecided, and others. We include students' self reported goals as an additional covariate for their postsecondary degree intentions. Lastly, we add additional controls for college-level means of our individual characteristics (eleventh grade CST math and English scores, race/ethnicity, gender, parental education, API, and student goal). Table 1 includes descriptive statistics on all of our measures at the individual level and Table 2 includes descriptive statistics at the college level.[13]

Empirical Methods

We begin by examining our outcomes across the community colleges in our sample. Figure 1 presents the distribution of Total Transfer Units, Proportion Persisting to Year 2, Proportion Transfer, and Proportion Completing across our 108 community colleges. To motivate the importance of accounting for student inputs, we plot each outcome against students' eleventh grade math test scores at the college level (Figure 2). From this simple scatterplot it is clear that average higher student test scores are

[13] Unlike the four-year college quality literature, we do not account for students' college choice set since most community college students enroll in the school closest to where they attended high school. Using nationally representative data, Stange (2012) finds that in contrast to four-year college students, community college students do not appear to travel farther in search of higher quality campuses, and, importantly, "conditional on attending a school other than the closest one, there does not appear to be a relationship between student characteristics, school characteristics, and distance traveled among community college students," (Stange 2012: p. 81).

associated with better average college outcomes. However, we also note considerable variation in average outcomes for students with similar high school test scores.

To examine whether there are significant differences in quality across community college campuses, we estimate the following linear random effects model:

$$Y_{isct} = \beta_0 + \beta_1 x_{isct} + \beta_2 \bar{x}_c + \beta_3 w_{is} + \zeta_c + \varepsilon_{isct}$$

where Y_{isct} is our outcome variable of interest (transfer units earned, persistence into year two, transfer to a four-year institutions, or degree/certificate completion for individual *i*, from high school *s*, who is a first time freshman enrolled at community college *c*, in year *t*; x_{isct} is a vector of individual-level characteristics (race/ethnicity, gender, parental education, and eleventh grade math and English language arts test scores), $\bar{x}_c$ are community college means of x_{isct}, and w_{is} is a measure of the quality of the high school (California's Academic Performance Index score)[14] attended for each individual. And ε_{isct} is the individual-level error term.

The main parameter of interest is the community college random effect, ζ_c.[15] We estimate $\widehat{\zeta_c}$ using an empirical Bayes shrinkage estimator to adjust for reliability. The empirical Bayes estimates are best linear unbiased predictors (BLUPs) of each community college's random effect (i.e. "quality"), which takes into account the variance (signal to noise) and the number of observations (i.e., students) at each college campus. Estimates of ζ_c with a higher variance and a fewer number of observations are shrunk towards zero (Rabe-Hesketh and Skrondal 2008).

[14] The Academic Performance Index (API) is a measure of California schools' academic performance and growth. It is the chief component of California's Public Schools Accountability Act, passed in 1999. API is comprised of schools' state standardized test scores and results on the California High School Exit Exam; scores range from a low of 200 to a high of 1000.

[15] We use a random effects model instead of fixed effects model due to the efficiency (minimum variance) of the random effects model. However, our findings are qualitatively similar when using a fixed effects framework.

The empirical Bayes technique is commonly used in measuring the quality of hospitals (Dimick, Staiger and Birkmeyer 2010), schools/neighborhoods (Altonji and Mansfield 2014), and teachers (Kane and Staiger 2008; Carrell and West 2010). In particular, we employ methodologies similar to those recently used in the literature to rank hospital quality, which shows the importance of adjusting mortality rates for patient risk (Parker et al. 2006) and statistical reliability (i.e, caseload size) (Dimick, Staiger, and Burkmeir 2019).   In our context, we similarly adjust our college rankings for "student risk" (e.g., student preparation, quality, and unobserved determinants of selection) as well as potential noise in our estimates driven by differences in campus size/student population.

IV. RESULTS

Are there measured differences in college outcomes?

Since we are interested in knowing if there are difference in student outcomes across community college campuses, we start by examining whether there is significant variation in our estimates of $\widehat{\zeta_c}$'s for our various outcomes of interest. Table 3 presents results of the estimated variance, $\sigma^2_{\hat{\zeta}}$, in our college effects for various specifications of equation (1). High values of $\sigma^2_{\hat{\zeta}}$ indicate there is significant variation in student outcomes across community college campuses, while low values of $\sigma^2_{\hat{\zeta}}$ would indicate that there is little difference in student outcomes across campuses (i.e., no difference in college "quality").

In Row 1, we start with the most naïve estimates where we include only a year by semester indicator variable. We use these estimates as our baseline model for comparative purposes and consider this to be the upper bound of the campus effects. These unadjusted estimates are analogous to comparing means (adjusted for reliability) in student outcomes across campuses. Estimates of σ_{ζ}^{2} in Row 1 show there is considerable variation in mean outcomes across California's community college campuses.

For ease of interpretation we discuss these effects in standard deviation units. For our transfer units completed outcome in Column 1, the estimated variance in the college effect of 4.86 suggests that a one-standard deviation difference in campus quality is associated with an average difference of 2.18 transfer units completed in the first year for *each* student at that campus. Likewise, there is significant variation across campuses in our other three outcome measures. A one-standard deviation increase in campus quality is associated with a 6.3 percentage point increase in the probability of persisting to year two (σ_{ζ}^{2}= .0042), a 7.3 percentage point increase in the probability of transferring to a four-year college (σ_{ζ}^{2}= .0056), and a 7.3 percentage point increase in the probability of completion (σ_{ζ}^{2}= .0056).[16]

One potential concern is that our estimates of σ_{ζ}^{2} may be biased due to differences in student quality (e.g., aptitude, motivation, etc.) across campuses. That is, the mean differences in student outcomes across campuses that we measure in Row 1 may not be due to real differences college "quality", but rather due to differences (observable or unobservable) in student-level inputs (e.g., ability). To highlight this potential bias, Figure 2 shows considerable variation across campuses in our measures of student ability.

[16] Completion appears to be driven almost entirely by transfer; that is, few students who do not transfer appear to complete AA degrees, as such, these two outcomes are likely measuring close to the same thing.

The across campus standard deviation in 11th grade CST math and English scores is 0.25 and 0.27 standard deviation, respectively.

Therefore, in results shown in rows 2-5 of Table 3 we sequentially adjust our estimates of $\widehat{\zeta_c}$ for a host of student-level covariates. This procedure is analogous to the hospital quality literature that calculates "risk adjusted" mortality rates by controlling for patient observable characteristics (Dimick, Staiger and Birkmeyer 2010). Results in Row 2 control for eleventh grade math and English standardized test scores. Row 3 additionally controls for our vector of individual-level demographic characteristics (race/ethnicity, gender, and parental education level). Results in Row 4 add a measure of student motivation, which is an indicator for student's reported goal to transfer to a four-year college. Finally, in Row 5 we add a measure of the quality of the high school that each student attended, as measured by California's API score.

The pattern of results in Rows 2-5 suggests that controlling for differences in student level observable characteristics accounts for some, but not all of the differences in student outcomes across community colleges. Results for our transfer units earned outcome in Column 1 show that the estimated variance in the college effects shrinks by 37 percent when going from our basic model to the fully saturated model. Despite this decrease, there still remains considerable variation in our estimated college effects, with a one standard deviation increase in campus quality associated with a 1.73 increase in the average number of transfer units completed by each student ($\sigma_{\hat{\zeta}}^2$= 3.07).

Examining results for our other three outcomes of interest, we find that controlling for student-level covariates shrinks the estimated variance in college quality by 26 percent for our persistence outcome, 70 percent for our transfer outcome, and 60 percent for completion. Again, despite these rather large decreases in the variance of the estimated college effects, there remains considerable variation in student outcomes across campuses. A one-standard deviation increase in college quality is associated with a 0.053 increase in the probability of persisting (σ_{ζ}^2= .0031), a 0.039 increase in the probability of transferring (σ_{ζ}^2= .0017), and a 0.045 increase in the probability of completion (σ_{ζ}^2= .0022). Graphical representations of the BLUPs from Model 5 are presented in Figure 3.

Although the estimates shown in Row 5 control for a rich set of individual-level *observable* characteristics, there remains potential concern that our campus quality estimates may still be biased due to selection on *unobservables* that are correlated with college choice (Altonji, Elder, and Tabor 2005). To directly address this concern, recent work by Altonji and Masfield (2014) shows that controlling for group averages of observed individual-level characteristics adequately controls for selection on unobservables and provides a *lower bound* of the estimated variance in school quality effects.[17]

Therefore, in results shown in Row 6 we additionally control for group-level means of our individual characteristics (eleventh grade CST math and English scores, race/ethnicity, gender, parental education and API Score). We find that controlling for college-level covariates shrinks the estimated variance in college quality compared to the naïve model (Model 1) by 65 percent for transfer units, 72 percent for our persistence outcome, 84 percent for our transfer outcome, and 79 percent for completion. Model 5 remains our preferred specification, however, even in this highly specified model, it is

[17] Altonji and Mansfield (2014) show that, under reasonable assumptions, controlling for group means of individual-level characteristics "also controls for all of the across-group variation in the unobservable individual characteristics." This procedure provides a lower bound of the school quality effects because school quality is likely an unobservable that drives individual selection.

important to note that we still find considerable variation in student outcomes across community college campuses. These lower bound estimates indicate that a one standard deviation increase in campus quality is associated with a 1.29 (10.9-percent) increase in the number of transfer units completed ($\sigma^2_{\hat{\zeta}}$= 1.68), a .034 (12.6-percent) increase in the probability of persisting ($\sigma^2_{\hat{\zeta}}$= .0012), a .030 (3.8-percent) increase in the probability of transferring ($\sigma^2_{\hat{\zeta}}$= .0009), and a .033 (9.7-percent) increase in the probability of completion ($\sigma^2_{\hat{\zeta}}$= .0011).

Exploring Campus Ranking

Given recent proposals by the Obama Administration to create a college scorecard, it is particularly critical to determine how stable (or unstable) our college quality estimates, $\hat{\zeta}_c$, are across specifications with various control variables. On the one hand, if our naïve estimates in Row 1 result in a similar rank ordering of colleges as the fully saturated estimates in Rows 5 and 6, then scorecards based on unadjusted mean outcomes will provide meaningful information to prospective students. On the other hand, if the rank ordering of the estimated $\hat{\zeta}_c$,'s are unstable across specifications, then it is critical that college scorecards be adjusted for various student-level inputs.[18]

To help answer this question, we examine how the rank ordering of our college quality estimates change after controlling for our set of observable student characteristics. Figure 4 graphically presents the unadjusted and adjusted estimated college quality effects for our transfer unit outcome (our preferred specification Model 5 from Table 3).

[18] Both hospital rankings and teacher quality rankings have been shown to be sensitive to controlling for individual characteristics. See for example, Kane and Staiger (2008) and Dimick, Staiger and Birkmeyer (2010).

The green squares represent the unadjusted effects, while the blue dots represent the effects and 95% confidence intervals after adjusting for student-level covariates. This graph highlights two important findings: 1) schools at the very bottom and very top end of the quality distribution tend to stay at the bottom and top of the rankings, and 2) there is considerable movement up and down in the middle of the distribution. This result indicates that unadjusted mean outcomes may be valuable in predicting the very best and very worst colleges, but they likely do a poor job in predicting the variation in college quality in the middle of the distribution. The same pattern can be noted in the other outcomes not pictured.

To provide a more detailed look at how the rankings of college quality change when adjusting for student-level covariates, Figures 5 plots rank changes in transfer units in the first year by campus. This graph show that the rank ordering of campuses change considerably after controlling for covariates. The average campus changed plus or minus thirty ranks, with the largest positive change in rank of seventy-five and the largest drop in rank of negative forty-nine.

These results highlight the importance of controlling for student-level inputs when estimating college quality. They also throw caution to policy-makers who may be tempted to rank colleges based on unadjusted mean outcome measures such as graduation rates and/or post-graduation wages.

V. CONCLUSION

Understanding quality differences among educational institutions has been a preoccupation of both policy makers and social scientists for over half a century

(Coleman, et. al. 1966). It has been well established that individual ability and socioeconomic factors bear a stronger relation to academic achievement than where one goes to school. In fact, when these factors are statistically controlled for, it appears that differences between schools account for only a small fraction of differences in pupil achievement. Yet, the influence of institutional quality differences in the postsecondary setting, particularly at the less selective two-year sector, where the majority of Americans begin their postsecondary schooling, has rarely been explored.

To help fill this gap, we employ data from California's Community College System to examine whether there are significant differences in student outcomes across college campuses. Our results show there are considerable differences across campuses in both short-term and longer-term student outcomes. However, much of these differences are accounted for by student inputs, namely measured ability, demographic characteristics, college goals, and unobservables that drive college selection. Nevertheless, after controlling for these inputs, our results show that important differences between colleges remain. What is the marginal impact of being at a better quality college? Our lower bound estimates indicate that going from the 10^{th} to 90^{th} percentile of campus quality is associated with a 3.32 (34.3 percent) increase in student transfer units earned, a 0.07 (9.6 percent) increase in the probability of persisting, an 0.09 (40.7 percent) increase in the probability of transferring to a four-year college, and an 0.08 (27.1 percent) increase in the probability of completion.

A natural follow-up question is what observable institutional differences, if any, might be driving these effects? A close treatment of what might account for these institutional differences in our setting is beyond the scope of this paper. However, prior work has identified several characteristics that may be associated with student success, including peer quality, faculty quality, class size or faculty-student ratio, and a variety of measures for college costs (Long 2008; Calcagno, Bailey, Jenkins, Kienzl, Leinbach 2008; Bailey et al. 2006; Jacoby 2006).

Finally, we note that identifying institutional effects is not purely an academic exercise; given today's policy environment, practitioners and higher education leaders are looking to identify the conditions and characteristics of postsecondary institutions that lead to student success. Given the recent push by policymakers to provide college scorecards our analysis furthers that goal for a critical segment of higher education, public open-access community colleges, and the diverse students they serve. Our results show that college rankings based on unadjusted mean differences can be quite misleading. After adjusting for student-level differences across campus, the average school rank in our sample changed by plus or minus thirty ranks. As such, our results suggest that policymakers wishing to rank schools based on quality should adjust such rankings for differences across campuses in student-level inputs.

REFERENCES

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

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," Journal of Political Economy, University of Chicago Press, vol. 113(1), pages 151-184, February.

Altonji Joseph and Richard Mansfield. 2014. "Group-Average Observables as Controls for Sorting on Unobservables When Estimating Group Treatment Effects: the Case of School and Neighborhood Effects," Working Paper.

Anelli, Massimo, 2014. Returns to elite college education: a quasi-experimental analysis, Working paper.

Bailey, Thomas, Juan Carlos Calcagno, Davis Jenkins, Timothy Leinbach, and Gregory Kienzl.. 2006. "Is Student-Right-to-Know All You Should Know? An Analysis of Community College Graduation Rates." *Research in Higher Education* 47(5):491-519.

Black, Dan and Jeffrey Smith. 2006. "Estimating the Returns to College Quality with Multiple Proxies for Quality." *Journal of Labor Economics* 24:701-728.

Black, Dan and Jeffrey Smith. 2004. "How Robust is the Evidence on the Effects of College Quality? Evidence from Matching" *Journal of Econometrics* 121:99-124.

Borman, Geoffrey, and Maritza Dowling. 2010. Schools and Inequality: A Multilevel Analysis of Coleman's Equality of Educational Opportunity Data *Teachers College Record*, v112 (5): 1201-1246.

Bowen, William G., Matthew M. Chingos, and Michael McPherson. 2009. *Crossing the Finish Line*. Princeton: Princeton University Press.

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

Brint, Steve. & Jerome Karabel. 1989. *The Diverted Dream: Community Colleges and the Promise of Educational Opportunity in America, 1900-1985.* New York: 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? *Economics of Education Review*, 27: 632–645.

Carrell, Scott E. and James E. West. 2010. "Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors," Journal of Political Economy, University of Chicago Press, vol. 118(3), pages 409-432, 06.

Clotfelter, Charles T., Helen F. Ladd, Clara G. Muschkin, and Jacob L. Vigdor. 2013. “Success in Community College: Do Institutions Differ?,” *Research in Higher Education*, 54:805-824.

Dale, Stacy B. and Alan B. Krueger. 2002. "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables." *Quarterly Journal of Economics* 117:1491-1527.

Dale, Stacy B. and Alan B. Krueger. 2012. "Estimating the Return to College Selectivity over the Career Using Administrative Earning Data ". Princeton.

Dimick, Justin, Douglas Staiger, and John Birkmeyer. 2010. “Ranking Hospitals on Surgical Mortality: The Importance of Reliability Adjustment,” Health Services Research, 45(6).

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.

Goldhaber, Dan, Stephanie Liddle, Roddy Theobald, and Joe Walch. 2010. “Teacher Effectiveness and the Achievement of Washington’s Students in Mathematics,” Center for Education Data & Research, University of Washington, Working Paper 2010-06.

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

Hoekstra, Mark. 2009. "The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach." *Review of Economics and Statistics* 91:717-724.

Hoxby, Caroline. 2009. “The Changing Selectivity of American Colleges.” *Journal of Economic Perspectives*, 23(4):95-118.

Hussar, William J., and Tabitha M. Bailey. 2009. Projections of Education Statistics to 2018 (NCES 2009-062). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Jacoby, Daniel. 2006. Effects of Part-Time Faculty Employment on Community College Graduation Rates. *The Journal of Higher Education*, 77(6):1081-1103.

Kane, Thomas J., Jonah E. Rockoff, and Douglas O. Staiger. 2008. “What Does Certification Tell Us About Teacher Effectiveness? Evidence from New York City.” Economics of Education Review 27 (6):615 631.

Kane, Thomas J., and Cecilia Elena Rouse. 1999. The Community College: Educating Students at the Margin between College and Work. *Journal of Economic Perspectives*, 13(1): 63-84

Long, Mark. 2007. "College Quality and early Adult Outcomes" *Economics of Education Review* 27:588-602.

Monks, James. 2000. "The Returns to Individual and College Characteristics: Evidence from the National Longitudinal Survey of Youth. " *Economics of Education Review* 19:279-289.

Mullen, Ann L., Kimberly Goyette, and Joseph A. Soares. 2003. "Who Goes to Graduate School? Social and Academic Correlates of Educational Continuation after College." *Sociology of Education* 76:143-169.

Parker, J. P., Z. Li, C. L. Damberg, B. Danielsen, and D. M. Carlisle. 2006. ‘‘Administrative versus Clinical Data for Coronary Artery Bypass Graft Surgery Report Cards: The View from California.’’ *Medical Care*, 44 (7): 687–95.

Rabe-Hesketh, Sophia and Anders Skrondal. 2008. Multilevel and Longitudinal Modeling Using Stata. College Station, TX: Stata Press, second ed.

Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 2005. “Teachers, Schools and Academic Achievement.” *Econometrica*, 73(2): 417–458.

Rosenbaum, James. 2001. Beyond college for all: career paths for the forgotten half. Russell Sage Foundation.

Rumberger, Russell, and Gregory Palardy. 2005. “Does Segregation Still Matter? The Impact of Student Composition on Academic Achievement in High School.” Teachers College Record, 107(9):1999–2045.

Sengupta, Ria, and Christopher Jepsen.2006. California’s Community College Students. *California Counts: Population Trends and Profiles*, *8* (2). Public Policy Institute of California.

Shulock, Nancy, and Colleen Moore. 2007. *Rules of the Game: How State Policy Creates Barriers to Degree Completion and Impedes Student Success in the California Community Colleges*. Sacramento, CA: Institute for Higher Education and Leadership.

Small, Mario L. and Christopher Winship. 2007. "Black Students' Graduation from Elite Colleges: Institutional Characteristics and Between-Institution Differences." *Social Science Research* 36:1257-1275.)

Stange, Kevin. 2012. "Ability Sorting and the Importance of College Quality to Student Achievement: Evidence from Community Colleges," *Education Finance and Policy*, 7 (1): 74-105.

TABLES & FIGURES

Table 1. Sample Descriptive Statistics (n=254,865)

Variable	Mean	Std. Dev.	Min	Max
Outcomes				
Transfer Units in Year 1	11.88	9.61	0	60
Ever Transfer	0.27	0.44	0	1
Persist to Year 2	0.80	0.40	0	1
Complete Ever	0.34	0.47	0	1
Covariates				
English Test Score	333.65	55.70	150	600
Math Test Score	291.64	48.98	150	600
Asian	0.08	0.27	0	1
Pacific Islander	0.01	0.08	0	1
Filipino	0.05	0.21	0	1
Hispanic	0.39	0.49	0	1
Black	0.07	0.25	0	1
White	0.40	0.49	0	1
Did Not State Race	0.01	0.08	0	1
Multiple Marks	0.00	0.00	0	1
Female	0.50	0.50	0	1
Parents' Less than High School	0.15	0.36	0	1
Parents' High School Diploma	0.22	0.41	0	1
Parents' Some College	0.28	0.45	0	1
Parents' College Graduate	0.25	0.43	0	1
Parents' Did Not State	0.10	0.30	0	1
Cohort 2005	0.14	0.35	0	1
Cohort 2006	0.20	0.40	0	1
Cohort 2007	0.22	0.41	0	1
Cohort 2008	0.23	0.42	0	1
Cohort 2009	0.21	0.41	0	1
Fall	0.82	0.38	0	1
Summer	0.18	0.38	0	1
High School API	707.91	79.00	272	987
Goal-Transfer with AA	0.46	0.50	0	1
Goal-Transfer without AA	0.12	0.32	0	1
Goal-2-Year AA Degree	0.04	0.19	0	1
Goal-2-Year Vocational Degree	0.01	0.10	0	1
Goal-Vocational Certification	0.01	0.08	0	1
Goal-Undecided	0.14	0.34	0	1
Goal-Unreported Goal	0.13	0.33	0	1

Source: Author's Calculations, California Community College Chancellor's Office

Table 2. Sample Descriptive Statistics by College (n=108)

Variable	Mean	Std. Dev.	Min	Max
Outcomes				
Transfer Units in Year 1	11.44	2.44	4.96	17.39
Ever Transfer	0.25	0.08	0.06	0.43
Persist to Year 2	0.77	0.07	0.53	0.90
Complete Ever	0.33	0.08	0.09	0.52
Covariates				
English Test Score (std)	-0.05	0.27	-0.79	0.56
Math Test Score (std)	-0.04	0.25	-0.72	0.44
Transfer Units in Year 1	11.44	2.44	4.96	17.39
Ever Transfer	0.25	0.08	0.06	0.43
Persist to Year 2	0.77	0.07	0.53	0.90
Complete Ever	0.33	0.08	0.09	0.52
English Test Score (Std)	-0.05	0.27	-0.79	0.56
Math Test Score (Std)	-0.04	0.25	-0.72	0.44
Asian	0.07	0.07	0.00	0.37
Pacific Islander	0.01	0.01	0.00	0.05
Filipino	0.04	0.05	0.00	0.27
Hispanic	0.37	0.20	0.06	0.91
Black	0.08	0.11	0.01	0.69
White	0.41	0.22	0.01	0.85
Did Not State Race	0.01	0.01	0.00	0.05
Multiple Marks	0.00	0.00	0.00	0.00
Female	0.50	0.04	0.39	0.65
Parents Less than High School	0.16	0.10	0.01	0.48
Parents High School Diploma	0.22	0.05	0.10	0.37
Parents Some College	0.28	0.07	0.15	0.54
Parents College Graduate	0.24	0.07	0.05	0.41
Parent Did Not State	0.10	0.05	0.02	0.22
Cohort 2005	0.12	0.09	0.00	0.48
Cohort 2006	0.18	0.10	0.00	0.52
Cohort 2007	0.21	0.10	0.00	0.75
Cohort 2008	0.23	0.11	0.00	0.63
Cohort 2009	0.26	0.19	0.04	1.00
High School API	703.26	45.03	588.34	799.11
Goal-Transfer with AA	0.43	0.12	0.06	0.67
Goal-Transfer without AA	0.10	0.05	0.00	0.25
Goal-2-Year AA Degree	0.04	0.03	0.00	0.25
Goal-2-Year Vocational Degree	0.01	0.01	0.00	0.07
Goal-Vocational Certification	0.01	0.01	0.00	0.07
Goal-Undecided	0.15	0.07	0.00	0.33
Goal-Unreported Goal	0.12	0.16	0.00	0.84

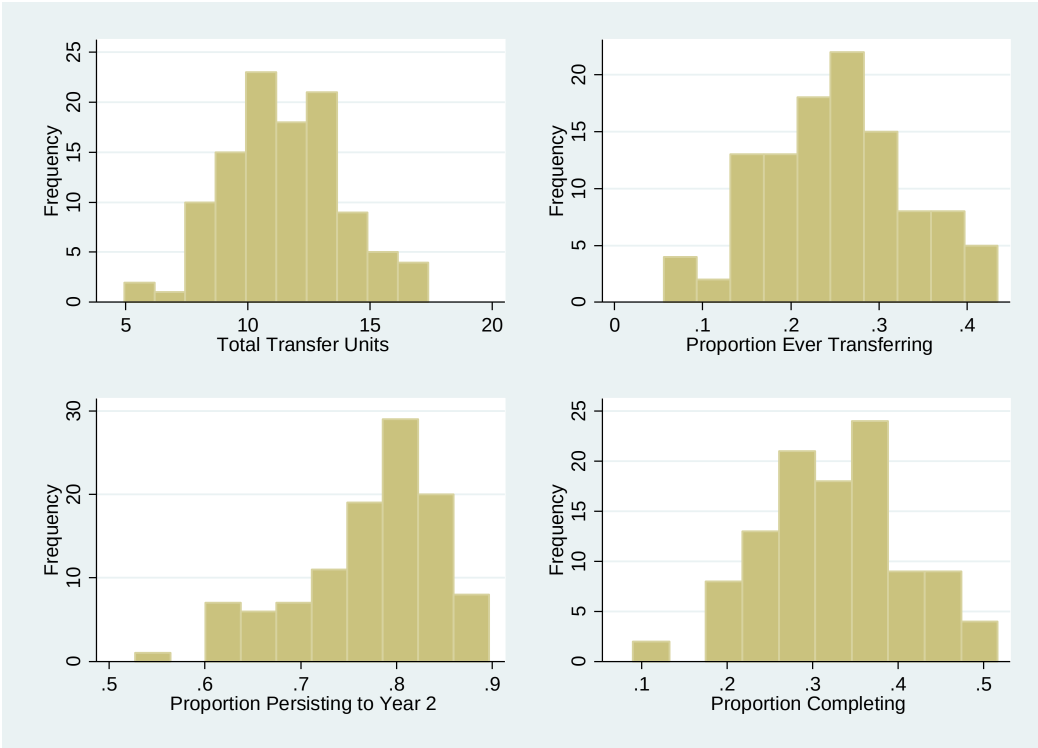
Source: Author’s Calculations, California Community College Chancellor’s Office

Table 3. Regression Results from Random Effects Models

Model		Transfer Units	Persist to Y2	Ever Transfer	Ever Complete
		Variance of Random Effects Estimates			
M1	Year/Term	4.86	0.0042	0.0056	0.0056
M2	Test Scores	3.69	0.0040	0.0034	0.0035
M3	Demographics	3.46	0.0038	0.0025	0.0029
M4	Goal	3.09	0.0032	0.0021	0.0025
M5	School API	3.07	0.0031	0.0017	0.0022
M6	College Means	1.68	0.0012	0.0009	0.0011
	% Variance Reduced M1 to M5	37%	26%	70%	60%
	% Variance Reduced M1 to M6	65%	72%	84%	79%

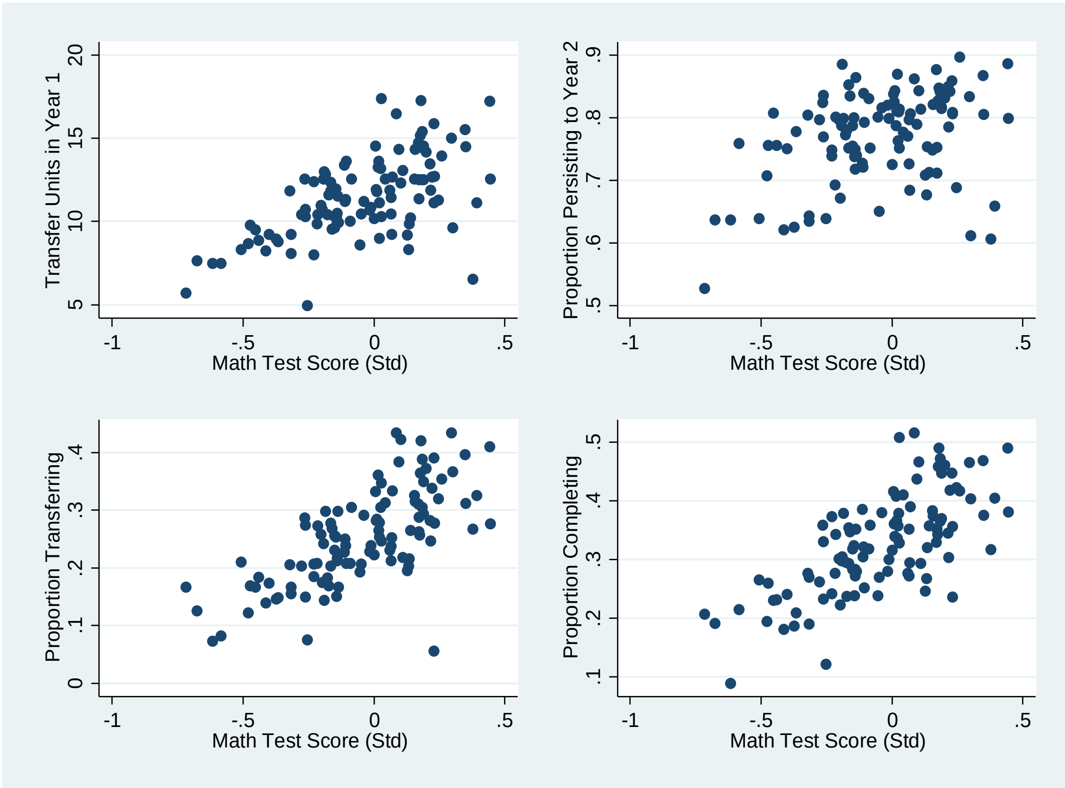
Source: Author's Calculations, California Community College Chancellor's Office

Figure 1. Distribution of Outcomes by College



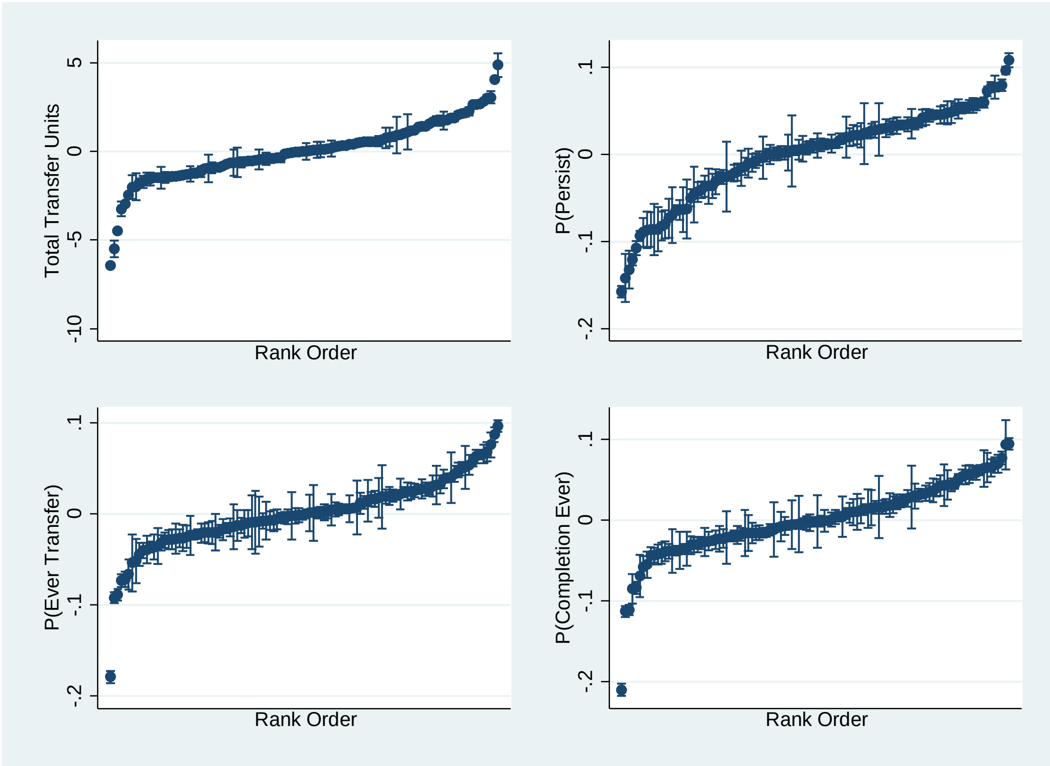
Source: Author's Calculations, California Community College Chancellor's Office

Figure 2. Scatterplot of Average College Outcomes against Students' 11th Grade Math Test Scores



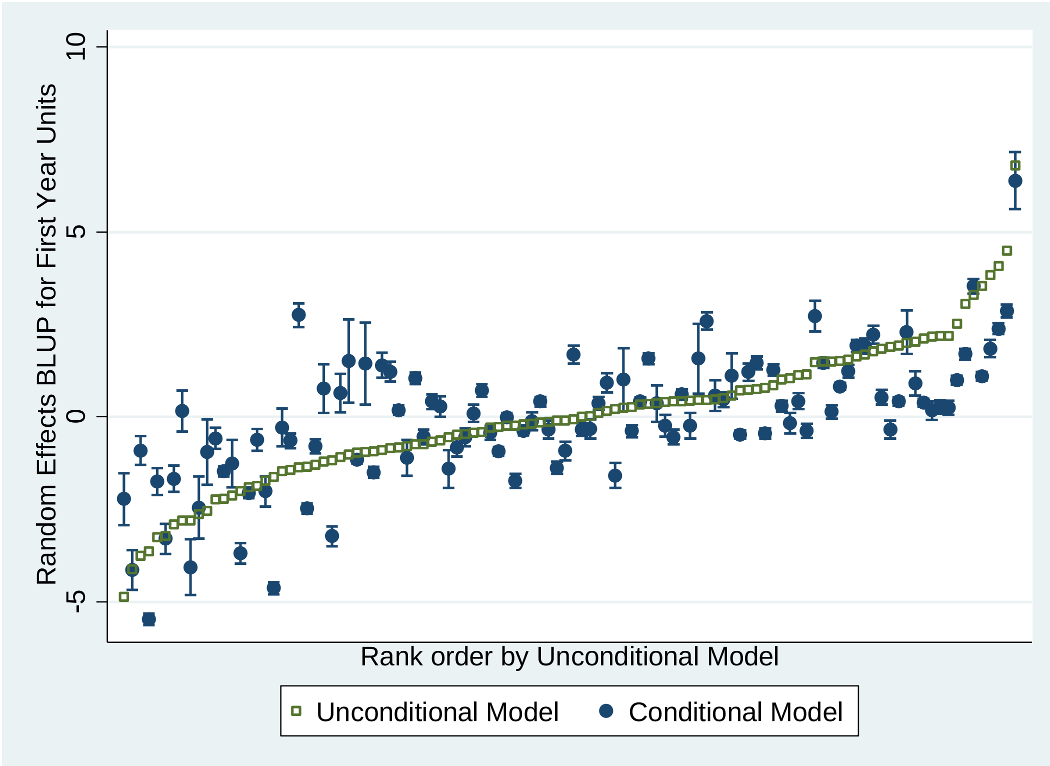
Source: Author's Calculations, California Community College Chancellor's Office

Figure 3. Ranked college effects for each community college, by outcome from fully specified model (Model 5, Table 3)



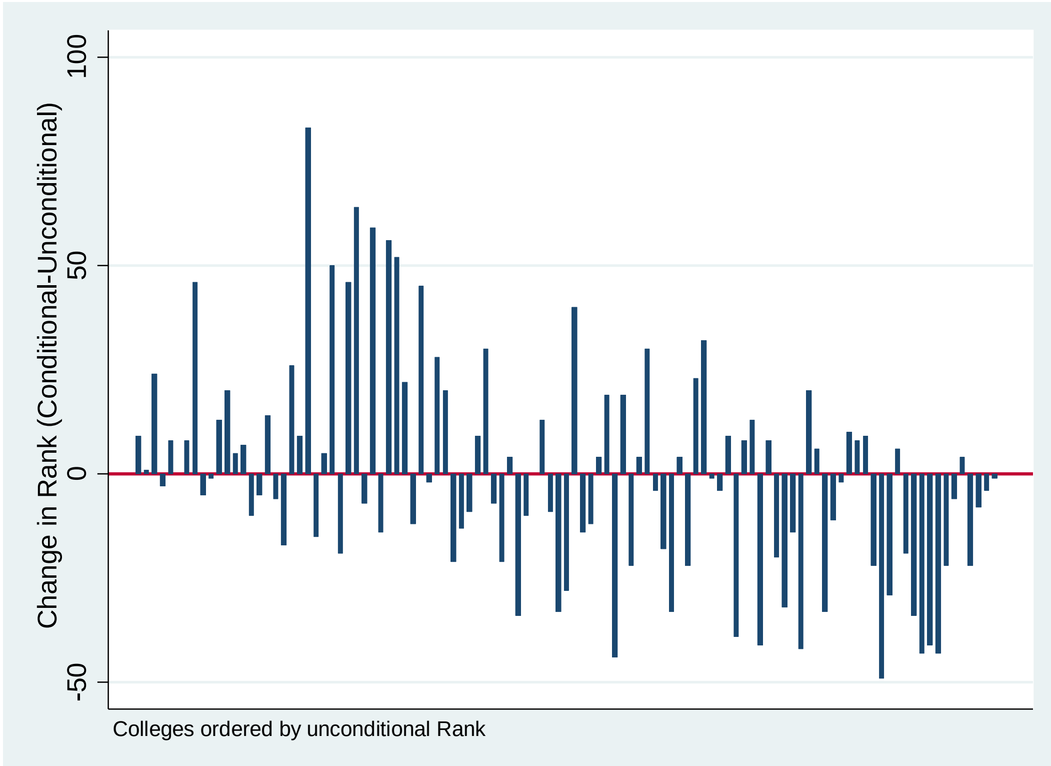
Source: Author's Calculations, California Community College Chancellor's Office

Figure 4. Unadjusted college effects compared to adjusted college effects for transfer units in the first year



Source: Author's Calculations, California Community College Chancellor's Office

Figure 5. Change in rank for each college from unadjusted model to fully specified model



Source: Author's Calculations, California Community College Chancellor's Office