CEDR Working Paper | WP 2014-7

How Much of A "Running Start" Do Dual Enrollment Programs Provide Students?

James Cowan and Dan Goldhaber
Center for Education Data & Research

University of Washington Bothell

The Center for Education Data & Research
University of Washington Bothell
3876 Bridge Way N. Ste. 201 Seattle, WA 98103
(206) 547-5585 cedr@u.washington.edu
www.cedr.us

Abstract. We study a popular dual enrollment program in Washington State, "Running Start" using a new administrative database that links high school and postsecondary data. Conditional on prior high school performance, we find that students participating in Running Start are more likely to attend any college but less likely to attend four-year colleges in the year after high school graduation. Additionally, we find evidence that data limitations common to previous studies of dual enrollment have substantial impacts on the estimates of dual enrollment on college outcomes.

Acknowledgements. The research presented here utilizes confidential data from the Education Research and Data Center located within the Washington Office of Financial Management (OFM). The views expressed here are those of the authors and do not necessarily represent those of the OFM or other data contributors. We gratefully acknowledge Shanna Jaggars, Eric Isenberg, Melissa Beard, Katie Weaver-Randall, Vivien Chen, Gary Benson, Jim Schmidt, David Prince, and participants at the Winter 2013 CALDER and Spring 2013 AEFP Conferences for helpful comments. We additionally thank ERDC for their expertise and support with the data used in this study. All errors are attributable to the authors.

Suggested citation:
Cowan, J. and Goldhaber, D. (2014). Title. CEDR Working Paper 2014-7. University of Washington, Seattle, WA.

Introduction

With more than 2 million students participating annually, dual-credit programs, which allow students to earn college credits while still enrolled in high school, have become the second most popular college preparatory program nationally after Advanced Placement. During the 2010-2011 school year, more than half of these students participated in programs that offered dual-credit courses on the campuses of postsecondary institutions (Thomas, Marken, Gray, Lewis, and Ralph, 2013). In the same year, 43% of public high schools nationwide offered students the opportunity to enroll in courses at a postsecondary institution. Because students earn college and high school credit simultaneously, these programs are seen as cost-effective way of increasing college readiness and college enrollment among high school students (An, 2013b; Bailey and Karp, 2013; State Board for Community and Technical Colleges, 2011). Yet, despite their popularity, there is relatively little empirical evidence of the impact of these programs on students' postsecondary enrollment choices or success in college. We use statewide data from Washington to investigate the effects of Running Start, Washington's dual enrollment program, on high school graduation and college enrollment.

We identify the graduation and enrollment effects of dual enrollment by controlling for student demographic variables and success in high school. Based on linear probability models, we find that dual enrollment students are more likely to earn an associate degree or enroll in college in the year following high school. However, we find no evidence that participation in dual enrollment increases full-time college attendance and some evidence that participation raises enrollment at two-year colleges at the expense of enrollment in four-year colleges.

While our identification strategy relies on strong assumptions about program participants, the longitudinal dataset we employ in this study nonetheless allows us to make two important contributions to the literature on dual enrollment. First, our sample identifies students before their participation in dual enrollment programs. This allows us to condition on measures of academic ability taken before participation in any dual enrollment programs and to capture outcomes for dual enrollment students who do not complete high school. We show that these restrictions influence our results, as dual enrollment students appear to have lower final grades than expected, perhaps due to stricter grading standards in college-level courses, and are less likely than non-participants to complete high school. Second, the data we analyze includes administrative records from the state higher education reporting system as well as college enrollment outcomes derived from the National Student Clearinghouse. We therefore track postsecondary enrollment with much greater accuracy than has been possible in previous research. Omitting postsecondary enrollments outside the public university system, which has been a common data limitation of prior studies, increases our estimates of the college-going effects of dual enrollment programs by 8-10 percentage points. In particular, the current study illustrates the value of integrating secondary and postsecondary datasets for researching high school programs that influence college outcomes.

Dual enrollment and educational attainment: An overview

Although the design of dual enrollment programs varies somewhat from state to state, three common features of these programs are particularly salient to students' educational attainment. First, dual enrollment programs aim to increase the rigor of the high school curriculum either by offering advanced courses at the high school or allowing students to enroll in courses on nearby college campuses. Second, they subsidize the costs of college by providing tuition-free enrollment in courses that count for college credit. Finally, they tend to have institutional features that may incentivize enrollment in particular kinds of colleges by locating courses on a college campus or by negotiating transfer agreements for credits earned while participating.

To the extent that dual enrollment programs provide a more rigorous high school curriculum than would be otherwise available, participation would be expected to improve the preparedness of students for college and facilitate more informed decisions about college selection. Most studies find that students who take a more rigorous high school curriculum have higher educational attainment and earnings (Adelman, 1999, 2006; Allensworth et al., 2009; Attewell and Domina, 2008; Aughinbaugh, 2012; Dougherty, Mellor, and Jian, 2006; Long et al., 2012; Morgan and Klaric, 2007; Rose and Betts, 2004; Speroni, 2011b), though these should not necessarily be interpreted as showing a causal relationship between high school curriculum and later outcomes as there exists the possibility that unobserved factors are correlated with high school attendance and educational attainment.[1]

The literature on the college dropout decision also suggests connections between the rigor of the high school curriculum and educational attainment. College students appear to respond to new information about their college-specific ability when deciding to continue investing in education (Arcidiacono, 2004; Stange, 2012; Stinebrickner and Stinebrickner, 2012). Dual enrollment programs may play a similar role for some students by providing them with low-cost information about their ability to succeed in college before making the more costly decision to enroll full-time. Participation in dual enrollment may therefore lead to better matches between student and college. Previous research has highlighted the important role of the quality of student-institutional matches in postsecondary persistence and completions (Arcidiacono, 2004; Light and Strayer, 2000). Consequently, dual enrollment policies may increase overall educational

[1] There is some quasi-experimental evidence on curriculum effects. Joensen and Nielsen (2009), for instance, examine a policy that increased the accessibility of advanced math courses and estimate an earnings premium of participation of about 20%, with the effect operating mainly through the increased likelihood of earning a postsecondary degree. Berger at al. (2014) study early college high schools using admissions lottery results as an instrument and find that students in such high schools are about 3 percentage points more likely to attend college after high school completion.

attainment by raising the graduation rate of those students who choose to enroll in college. For instance, An (2013a) finds that dual enrollment students who enroll in college are more likely to complete a degree than non-participants.

One of the touted benefits of dual enrollment programs is that they reduce the financial cost to students of obtaining college credits. The dual enrollment program we study in Washington covers up to the $4,000 full-time tuition at state two-year colleges, which amounts to an aggregate annual tuition subsidy of $39.7 million or about $2,000 per student per year (State Board of Community and Technical Colleges, 2006, 2010a). Additionally, by accumulating college credits while in high school, students may reduce the time it takes to complete a college degree, thereby reducing the opportunity costs of completing college.[2] By reducing the cost of obtaining a postsecondary degree, dual enrollment programs should unambiguously increase the likelihood that students enroll in some type of college after high school (Dynarski, 2003; Kane, 2007; van der Klaauw, 2002). However, as noted by Manski (1989), such policies might have a larger enrollment than completion effect if they induce more students with a higher probability of dropout to enroll.

Institutional features of dual enrollment programs may also influence the college enrollment decisions of participants. The Washington State program is housed at state community colleges. If college completion outcomes are uncertain, students who have nearly completed an associate degree may choose to enroll in a two-year college to take advantage of any sheepskin effects of the two-year degree (Light and Strayer, 2004). Moreover, the community colleges have transfer agreements with the in-state public, four-year universities that facilitate the transfer of credits and completion of major requirements. Dual enrollment participation may therefore lead some students to substitute preliminary coursework at a community college for preliminary coursework at a four-year institution. The empirical literature has reached divergent conclusions about how this could affect the probability of earning a four-year degree. On the one hand, this initial diversion could lead students to be better prepared for the rigor of four-year colleges and provide an opportunity to transfer to a higher quality college (Hilmer, 1997). On the other hand, several suggest that students with who first enroll in a two-year college with plans to transfer are less likely to complete a bachelor's degree than first-time enrollees at four-year colleges (Leigh and Gill, 2003; Long and Kurlaender, 2008; Rouse, 1995), although there is also some evidence that the "diversionary" effect of two-year colleges reflects differences in educational plans (Leigh and Gill, 2003). In summary, it appears from existing empirical work on community colleges that dual enrollment programs might be expected to increase postsecondary enrollment but that part of this increase may be offset by a reduction in the probability of students'

[2] In the Washington State program, the average student takes 11 credits per quarter, or more than 60 credits if enrolled for both junior and senior years (State Board for Community and Technical Colleges, 2011).

completing bachelor's degrees conditional on enrollment. The overall effect on educational attainment is therefore unclear.

Review of the dual enrollment literature

Despite their popularity, there is relatively little evidence on the effects of dual enrollment programs on college attendance or completion. This reflects the difficulty in collecting information on student's high school academic history and college enrollment patterns. Until recently, statewide databases have not typically linked the academic records of high school students to postsecondary outcomes. Given this the majority of previous studies have suffered from one of two missing data problems. First, several studies rely on selected samples of students where sample selection may be influenced by dual enrollment participation. Second, the datasets employed in most studies of college enrollment lack college enrollment data for a substantial portion of the sample.

The literature on the college completion effects of dual enrollment largely estimates effects for students who enroll in college, an outcome that may itself be influenced by participation. In a particularly careful analysis, An (2013a) analyzes data from the NELS:88 and finds that dual enrollment participants who enroll in college are about 8 percentage points more likely to complete any college degree and 7 percentage points more likely to complete a bachelor's degree than non-participants. Other studies have examined the outcomes of students who enroll in particular postsecondary institutions and have found that students with prior dual enrollment participation generally earn higher grades and persist at greater rates (Allen and Dadgar, 2012; Karp et al., 2007).

While informative, estimates of program effects using selected samples represent the sum of two separate contrasts. The first is the causal effect of dual enrollment on college completion for students who participate in dual enrollment and register for college. This provides an estimate of the treatment effect for a subset of the all students who participate in dual enrollment programs. The second portion of the estimated effect is the difference in counterfactual completion rates for students who would attend college with dual enrollment participation and those who would attend college without participation.[3] This reflects the fact that dual enrollment affects the likelihood of completing college and the composition of those who enroll (Angrist, 2001). If dual enrollment programs uniformly increase college enrollment, this second effect may be negative and the estimated effect could understate the causal effect of dual enrollment on

[3] Let R represent college enrollment, C represent college completion, and D indicate dual enrollment participation. Further, let subscripts (0, 1) indicate the possibly counterfactual outcome that would be observed for students with D=0 and D=1, respectively. Then, if we assume conditional independence of (C, R) and D given covariates X, we can write the estimated effect among the college enrolled sample as $E[C|D=1, R=1, X] - E[C|D=0, R=1, X] = E[C_1 - C_0|D=1, R_1=1, X] + (E[C_0|D=1, R_1=1, X] - E[C_0|D=1, R_0=1, X])$.
The first term is the causal effect of interest, while the second is the result of selection into the sample.

college completion (Manski, 1989). On the other hand, if dual enrollment programs provide students with information about their ability to succeed in college and lead to better college-going decisions, estimated effects may actually overstate the causal college completion effect. Evidence on the college enrollment effect of dual enrollment programs may therefore be important for assessing the available long-term evidence on college completion.

Previous studies of the college transitions of dual enrollment students have estimated a wide range of effects on college enrollment. Karp et al. (2007) and Struhl and Vargas (2012) estimate that dual enrollment programs in Florida and Texas, respectively, have substantial effects on college enrollment. Karp et al. (2007), for instance, estimate that participation in Florida's dual enrollment program increases the probability of any college enrollment by 16.8 percentage points and of university enrollment by 7.7 percentage points. Struhl and Vargas (2012) estimate an effect on the odds ratio of college attendance of 2.3.[4] Speroni (2011b) compares dual enrollment participation to AP participation and finds a relative effect of about 6 percent in favor of dual enrollment. Using a regression discontinuity design that exploits a minimum GPA eligibility requirement for some of Florida's dual enrollment programs, Speroni (2011a) finds no statistically significant effect of participation on either postsecondary enrollment or completion. The enrollment effects are imprecisely estimated, however, and 95% confidence intervals cover the effects found in the previous literature.

One limitation of many prior studies is the reliance on state data warehouses that omit records on private or out-of-state enrollments. The dataset used by Karp et al. (2007) includes college enrollment records only for students at in-state, public institutions. Similarly, the dataset used by Struhl and Vargas (2012) includes only enrollment in public or private institutions in Texas. These omissions may lead to substantial misclassification rates. Notably, Speroni (2011a, 2011b), which find smaller or null effects, rely on samples where data on in-state enrollments is combined with data from the National Student Clearinghouse. These omissions may lead to substantial classification error in the outcome variables of interest. For instance, Speroni (2011a) finds that in the dataset analyzed by Karp et al. (2007), 12% of all students are incorrectly classified as not attending college. Unlike with continuous dependent variables with classical measurement error, this form of classification error generally biases coefficient estimates. If the probability of misclassification is constant across observations, misclassification will tend to attenuate coefficient estimates (Hausman et al., 1998; Meyer and Mittag, 2013). However, in practice, the bias will depend on how the probability of misclassification varies with the sample covariates. We discuss this issue in more detail below. In our dataset, about 17% of students enroll in a private or out-of-state

[4] Struhl and Vargas (2012) do not report college enrollment rates, but if the enrollment rates of dual enrollment students in Texas are similar to those in Washington and Florida, the estimated odds ratios imply marginal effects in the range of approximately 0.13-0.19.

institution and we find that misclassifying these students as not attending college generates an upward bias in estimates of college enrollment effects of about 8-10 percentage points.

Washington State's Running Start program

Washington's dual enrollment program, Running Start, started statewide in 1992 and has enrolled more than 10% of the state's high school juniors and seniors since the 2007-2008 school year (State Board of Community and Technical Colleges, 2009). As with many other dual enrollment programs, Running Start allows juniors and seniors to take courses tuition-free at any of the state's 34 community colleges.[5] In Washington State, community colleges alone determine eligibility, which typically requires placement into a college-level English or mathematics course using a placement exam such as COMPASS or Accuplacer. Washington law specifically prohibits high schools from conditioning participation on administrator approval or high school academic record. This arrangement in Washington is far less restrictive than the norm for dual enrollment programs: 77% of schools nationwide require the permission of a counselor or administrator and 49% require a minimum cumulative grade point average (Thomas et al., 2013).[6]

School districts pay the community college 93% of the state basic education allotment for each full-time student participating in the program. The state estimates the total tuition subsidy cost $41.3 million for the 2009-2010 school year (State Board of Community and Technical Colleges, 2011). As of 2010, colleges received $4,500 per full-time equivalent Running Start student, which is estimated to represent only 60% of the cost of educating a two-year college student (State Board of Community and Technical Colleges, 2010b, 2011).

Once students enroll in Running Start, they may take a combination of high school and college courses. Although all community colleges offer some distance education (mostly online) and many Running Start students take courses online, less than 1% of Running Start students take all their courses online.[7] Again, this is in marked contrast to the national norm for such programs: nationally, only 43% of schools with participants in academic-oriented dual enrollment programs have students attending

---

[5] Three universities also participate in the program, however we focus here on Running Start participation at the community college level because we have incomplete data on participants at four-year colleges and it is difficult to distinguish the four-year college participants from students taking other transitional courses, such as College in the High School programs, in which students earn credit from courses taught on high school campuses by high school faculty members. More than 98% of students participate at a community or technical college. During the years we consider, 300-400 students participated each year at four-year colleges, while more than 16,000 participated each year at two-year colleges (State Board of Community and Technical Colleges, 2006, 2008, 2009, 2010a).

[6] Since 2010, Running Start has required students to meet with a counselor to complete their registration. However, high schools may not condition participation on the approval of a counselor.

[7] 10% of Running Start students take a majority of their courses online.

school on a postsecondary campus (Thomas et al., 2013). Given the differences in eligibility requirements and the location of services, students in Washington's Running Start program appear to have much greater independence than students in similar programs in other states.

Although students may elect to attend college full-time while enrolled in Running Start, they must complete their district's graduation requirements to receive a high school diploma. School districts generally map specific courses taken at community colleges to particular state and district graduation requirements.[8] Consequently, more than half of Running Start students take credits amounting to full-time enrollment at a community college (State Board of Community and Technical Colleges, 2011). In our sample, Running Start students attempt an average of 47 credits, or more than one academic year, while in high school, and 13% of participants actually earn an associate degree by high school graduation.

Upon graduation from high school, Running Start students have the same options for college enrollment as traditional high school graduates. Many, for instance, continue in the community college system and earn an associate's degree. In our sample, in fact, about 30% of Running Start students continue at the *same* college immediately after graduation. Although they may have a substantial number of credits, the in-state public universities treat Running Start students who have not completed an associate degree as freshmen for admissions purposes. Alternatively, if they choose to complete an associate degree, Running Start students may apply to public four-year colleges as transfer students.

Given the myriad college pathways open to participants, Running Start students may have varied college plans at the time of enrollment. Using data collected during the registration process, we estimate that 24% of Running Start students plan to complete an associate degree and transfer to a four-year college, 41% have no plans for an associate degree but do plan to attend a four-year college, 8% plan to complete an associate degree only, and 27% have no specific plans for college.[9] Following high school, 73% of

[8] Students who complete an associate degree before graduation can also apply for a high school completion certificate from their college.

[9] We construct college enrollment intent using two questions asked of incoming community college students. The first asks about students' planned length of attendance at the community college and the second about the purpose of enrolling. We use the following classification:
AA and BA: Purpose is "Transfer to a four-year college" and planned attendance is "Long enough to complete a degree"
BA only: Purpose is "Transfer to a four-year college" and planned attendance is "One quarter", "Two quarters", "1 year", or "Up to 2 years, no degree planned"
AA only: Purpose is "Take courses related to current or future work", "High school diploma or GED certificate", "Explore career direction", or "Personal enrichment" and planned attendance is "Long enough to complete a degree"
No degree: Purpose is "Take courses related to current or future work", "High school diploma or GED certificate", "Explore career direction", or "Personal enrichment" and planned attendance is "One quarter", "Two quarters", "1 year", or "Up to 2 years, no degree planned."

Running Start students enroll in any college and 36% enroll full-time in a four-year college.

Data

We use data on high school and college students from the Education Research and Data Center (ERDC) warehouse in Washington State. The ERDC data include high school enrollment and standardized test records from the Washington Office of the Superintendent of Public Instruction (OSPI); community college attendance, transcript, and degree completion data from the Washington State Board of Community and Technical Colleges (SBCTC); and university attendance, transcript, and degree completion data from the Public Higher Education Enrollment System (PCHEES). Data on current and former Running Start participation are included in the SBCTC and PCHEES data systems. In addition to data on enrollments in public colleges in the state, the ERDC data includes information on enrollment in private and out-of-state enrollments from the National Student Clearinghouse (NSC).

We focus on the cohort of high school students who first enrolled in ninth grade during the 2006-2007 school year. Our analysis dataset contains all students who enrolled in a standard high school during their sophomore year and continued their enrollment in the fall of their junior year.[10] As students have the option of enrolling in Running Start during the 11th and 12th grades, we study the outcomes for students participating in Running Start for school years 2008-2009 and 2009-2010. Our analytic sample includes 55,396 students, of whom 8,586 participated in Running Start.[11] Table 1 provides mean student characteristics for students by their Running Start participation status.

We construct high school graduation measures using records from OSPI and SBCTC. For all high school completion variables, we include only outcomes that occur within four years of student's initial high school enrollment; that is, we consider completions that occur by the summer of 2010. Therefore, our outcomes properly measure on-time high school completion. The OSPI student reporting system explicitly codes students confirmed to have dropped out of school. Our dropout variable is defined using this code and may exclude students whose status is not confirmed by the reporting school. We obtain records on GED attainment from both OSPI, which records receipt of the GED as a possible graduation outcome, and SBCTC, which administers the test. Because Running Start students may obtain a high school diploma equivalent from their

---

It is important to note that the individual questions have not been validated for Running Start. Construction of the derived variables and the assumptions that underlie them are purely the work of the authors (David Prince, personal communication, June 19, 2014).

[10] We exclude students enrolled in alternative high schools, some of which participate in Running Start or other dual enrollment programs, from this analysis. We also exclude private school and home school students who enroll part-time in a public high school.

[11] This represents 78% of the total number of students enrolled statewide in all types of high schools during this period.

community college for either completion of the associate degree or of a credit-based high school equivalency program, we include any credit-based form of high school completion in our high school diploma measure. This definition is consistent with prior research that has found a distinction between credentials awarded for credit and those awarded for passing a test (Cameron and Heckman, 1993). When constructing the GED and dropout measures, we exclude students who otherwise obtain a valid high school diploma.

We obtain measures of student postsecondary enrollment from the SBCTC, PCHEES, and NSC databases. For the in-state public institutions, we derive college enrollment based on whether students register for a positive number of credits during the fall term of 2011 (the first term after scheduled graduation). Similarly, for students in the NSC database, we define college enrollment using the enrollment status reported by the NSC. We additionally consider whether students are enrolled full-time in college. For the in-state schools, we define full-time enrollment as attempting at least 12 credits for institutions on the quarter system and 10 credits for institutions on the semester system. As Running Start students may earn an associate degree during high school, the measures of college enrollment we use in our regressions include students who have enrolled in a two-year or four-year college or have completed an associate degree. We additionally identify students whose initial enrollment is in a four-year college following high school graduation.

We obtain pre-treatment student academic achievement and program participation information from OSPI enrollment records. We measure all student characteristics at the end of the sophomore year to avoid controlling for factors, such as grades, special education status, or free or reduced price lunch eligibility, which may be endogenous to Running Start participation. All tenth graders in the state take a standardized exam in reading, mathematics, and writing that we use in this analysis. Data on student grades, demographics, and program participation are also available from OSPI.

As is apparent from Table 1, Running Start students are more likely to attend college after graduation than non-participants. Seventy-three percent of Running Start students attend any college in the year after they graduate and 59% do so full time. Both are far higher than the corresponding statewide means of 59% and 49%. However, Running Start students also have substantially better academic high school performance records than the overall population of high school students. They score about 0.4 standard deviations higher on the tenth grade standardized tests than the sample average and they have an average tenth grade GPA (3.23) that is nearly 0.5 points higher than the state average (2.86). They are also less likely to participate in special education, bilingual education, Title I, and/or the federal free and reduced price lunch program.

Dual enrollment, high school completion, and college attendance

We are concerned with estimation of the effect of Running Start participation on the probability of high school completion and college enrollment. As with much of the previous research on dual enrollment, we employ a selection on observables design to estimate the effects of participation.[12] That is, we estimate

$$Y_i = X_i\beta + \delta RS_i + \epsilon_i, \tag{1}$$

where Y denotes high school completion or college enrollment, X denotes the vector of observed student covariates, and RS denotes participation in Running Start. We begin by estimating linear probability models that condition on a number of student characteristics that may influence postsecondary enrollment decisions. These include a cubic polynomial in sophomore grade point average and test scores in mathematics, reading, and writing from a state standardized test administered at the end of tenth grade, student gender, ethnicity, free and reduced price lunch status, student learning disability status, participation in a bilingual education program, an indicator for a primary language other than English, participation in a targeted Title I program, participation in gifted and talented classes, AP classes, migrant status, and unexcused absences. All student controls are measured during the student's tenth grade year before eligibility for Running Start. In addition, we include school fixed effects in all regressions and cluster standard errors at the school level. The treatment effect δ is identified only under the assumption that unobserved factors associated with high school completion and college enrollment, ϵ_i, are uncorrelated with Running Start participation. This requires that students participating in a college preparatory program have similar college plans as observationally similar non-participants, an assumption that is unlikely to hold. We return to the plausibility of this assumption below.

The outcomes we consider are limited to short-term measures of educational attainment. In particular, we lack data on overall educational attainment. As discussed above, there are several reasons to believe that effects on initial enrollment and educational attainment will differ. Participation in dual enrollment programs may increase the likelihood of college graduation either directly by better preparing students for college-level courses or indirectly by inducing students with a greater likelihood of college completion to enroll. An (2013a, b) provide some evidence for this view. Moreover, dual enrollment programs that are located on community college campuses may encourage students to complete an associate degree at a two-year college before moving to a four-year institution. As noted above, among Running Start students in our dataset, 30% continue enrollment in the same college in the quarter after graduation. Conversely, the educational subsidy effects of dual enrollment may encourage more marginal students to enroll in college. Therefore, shot-term measures of enrollment in four-year colleges may understate the long-term effects of such programs on educational attainment. Nevertheless, even a focus on such short-term outcomes provides an important contribution to the dual enrollment literature. Prior research has relied mostly on samples with substantial measurement error in the postsecondary enrollment variables.

[12] Speroni (2011a), which uses a regression discontinuity design, is an exception.

We therefore provide some of the first evidence of such effects from a sample taken before participation and that includes a measure of postsecondary enrollment with little measurement error.

We report the OLS results in Panel A of **Table 2**. Compared to similar students, Running Start participants are less likely to earn a traditional high school diploma and more likely to drop out of school. Specifically, the point estimates suggest that Running Start students are 2.3 percentage points less likely to earn a credit-based diploma, 1.1 percentage points more likely to drop out of school and 0.4 percentage points more likely to earn a GED. All estimated coefficients are statistically significant at the 1% level. These findings may seem counterintuitive, but there are several potential explanations.

First, Running Start students may view an interim high school diploma as unnecessary given the possibility of earning a two-year college credential. Some Running Start participants may therefore continue their enrollment in community college without completing a high school degree. Among the Running Start students who fail to complete an on-time high school diploma, 15% are still enrolled in the college in which they participated in Running Start in the year after their scheduled graduation. In order to capture this possibility, we estimate the effect of Running Start participation on the probability that a student neither completes an on-time diploma nor enrolls in college. We find that Running Start students are 1.1 percentage points more likely to have done so. That is, apart from the increased risk of dropping out of school, it appears that Running Start students are more likely to either delay graduation or substitute the associate's degree for a formal high school credential.

Second, the relatively inclusive eligibility requirements of Running Start may attract a disproportionate share of non-traditional high school students who would be unlikely to complete a high school degree in the absence of the program.[13] If differences in the propensity to graduate high school are not captured by our measures of academic achievement, then the dropout effects may partially reflect selection into Running Start. On the other hand, even if the coefficients are unbiased, the eligibility rules may permit weaker students who may not benefit from college-level courses to participate in Running Start. We explore the latter possibility by analyzing the heterogeneity in Running Start effects by academic preparation. In Panels B, we estimate models that interact the Running Start indicator with a measure of academic achievement. To create the index, we regress the college enrollment indicator on tenth grade GPA and standardized test scores for students who do not enroll in Running Start and generate quintiles of the predicted values for each observation in the sample. The index is therefore a weighted mean of students' academic achievement variables where the weights reflect their importance for

[13] As we discussed earlier, Running Start does not condition participation on the approval of high school officials or have a statewide GPA requirement, although a few individual community colleges do. By contrast, 77% of schools with dual credit programs nationwide require administrator approval and 49% have minimum GPA requirements (Thomas et al., 2013).

predicting college attendance. When we disaggregate the Running Start effect in this way, we find that the high school graduation results are concentrated among students with low predicted probabilities of college enrollment. Students in the bottom quintile of academic preparation who participate in Running Start are 4.6 percentage points more likely to drop out and 4.1 percentage points more likely to earn a GED than similarly prepared non-participants. Similarly, we estimate that Running Start students in the bottom quintile are 4.8 percentage points less likely to earn a high school diploma, which is marginally statistically significant, and students in the second quintile are 5.8 percentage points less likely, which is statistically significant. We estimate statistically significant and negative graduation effects for the next two quintiles that are substantially smaller and find no effect on graduation for the most prepared Running Start students. Taken together, the results suggest that Running Start students are more likely to drop out of high school, but that estimates of on-time graduation effects may overstate the tendency for participants to fail to earn a secondary or postsecondary credential. Specifically, when we include college enrollees in the high school completion category, only about half of the negative completion effect remains.

In contrast to the negative high school completion findings, Running Start students are more likely to attend college or have earned a degree by the first term after graduation: Running Start students are 5.3 percentage points more likely to either attend any college or have already earned an associate degree, but they are no more likely to have either earned a degree by the fall term after high school or to attend college full-time. Furthermore, Running Start students are 9.1 percentage points *less* likely than similar non-participants to attend a four-year university full-time. In results not shown, we find that the negative effect on four-year enrollment is strongest for students who earn nearly, but not quite, enough credits through Running Start to obtain an associate degree. Depending on the rate at which these students transfer to four-year colleges, some of the negative effect on four-year college enrollment may be temporary.

When we allow the effects of Running Start to vary by measures of students' college preparation, we find that lower-performing students appear to gain the most from participation, which is consistent with previous studies of dual enrollment (Karp et al., 2007; Speroni, 2011b). The college enrollment effects among students in the bottom three quintiles of the academic achievement index distribution all exceed 6 percentage points and are statistically significant, while we observe small, but positive, effects for the top two quintiles. Similarly, we only observe positive full-time enrollment differences for students in the second quintile, whereas we estimate null or negative differences for all other participants. Finally, the estimates of the diversion effect of Running Start are small and not statistically significant at conventional levels for the least prepared students.

Our identification strategy assumes that a limited set of academic and financial information sufficiently control for differences in educational attainment between Running Start and non-Running Start students. However, as noted above, students may select into Running Start based on financial and academic factors that are unobserved. Because dual enrollment programs offer very clear benefits to students who intend to enroll in college after high school, we might expect estimation strategies that assume selection on observables to overstate the true effect of these programs on the probability of college enrollment. On the other hand, the literature on college dropouts suggests that the informational benefit of dual enrollment programs may be greatest for students with the most uncertainty about their ability to succeed in college (Stange, 2012). If this is the case, students who choose to participate may be otherwise less likely to attend and persist in college than observationally similar students. Similarly, estimated program effects that control only for student background may overstate the diversionary effect of dual enrollment programs. Because students can make considerable progress toward an associate degree while in high school, dual enrollment students may at baseline be more likely to plan to initially enroll in a two-year college.

We present estimates of the Running Start effect interacted with pre-participation college plans in Panel C. In column (4), we only find college enrollment effects for students who indicate that they plan to complete a college degree. Students who state no plan to complete a college degree are no more likely than similar students to enroll in college and less likely to do so full-time. In column (6), we find that the effects of Running Start on four-year enrollment are strongest for students who indicate intent to complete an associate degree. While we find an estimate of -0.136 for students who plan to complete an associate degree before transferring to a four-year institution, we find an effect of only -0.061 for students planning to enroll directly in a four-year institution. We interpret this as suggestive than selection into Running Start explains at least some of the estimated treatment effects. However, we note that this test must be taken with some caution. Running Start may increase educational attainment by increasing enrollment rates of students who consider themselves college bound rather than by increasing the college attendance rates of students who would otherwise not consider college. Similarly, Running Start students who declare their intention to earn an associate degree might have enrolled in a four-year college in the program's absence. Nonetheless, we do find some evidence that outcomes are correlated with college plans for Running Start students.

While the findings in Table 2 suggest that Running Start may shift some of the initial college enrollment toward two-year colleges, the results may not generalize to overall educational attainment. Among the 2011 graduating classes at Washington public universities, 40% of students had transferred from two-year colleges (Washington State Board for Community and Technical Colleges, 2013). Therefore, it is likely that some of these students will eventually transfer to four-year colleges. An (2013a) provides some support for this view, suggesting that dual enrollment students who enroll in any college are more likely to complete a bachelor's degree than similar non-participants. Hence, these results should be interpreted as a short-run effect and are not necessarily indicative of overall educational attainment.

Nonetheless, our results suggest something of a paradox for the benefits of dual enrollment programs. If our estimates represent a causal relationship, then we find that participation in Running Start has both relatively large college enrollment effects and relatively large high school dropout effects for students on the margin of college attendance. Moreover, the effect on high school non-completion persists even when we count students enrolled in community colleges as having obtained a high school credential. These findings suggest that policymakers considering eligibility policies for dual enrollment programs may need to balance increased access to college against the costs to students of failing to complete a high school degree.

The role of missing data in assessing dual enrollment programs

Our estimates of differences in college enrollment rates by dual enrollment participation are substantially smaller than those found in previous studies of statewide programs. While Karp et al. (2007) and Struhl and Vargas (2012) find college enrollment effects of participation in dual enrollment programs of about 15-20 percentage points, we find effects of only about 5 percentage points. One key difference between our study and these previous studies is that we have a measure of college attendance that includes enrollment at private or out-of-state institutions. We therefore consider the role that this data plays in our analysis.

The effect of misclassifying college attendance depends on the relationship between dual enrollment participation and the included covariates. In dual enrollment studies, the misclassification almost always consists of students who actually attend college but for whom enrollment records are missing. Typically, college students lack enrollment records because they attend private or out-of-state schools that are not included in state administrative databases. Of the several studies that consider the college attendance effects of dual enrollment described earlier, only Speroni (2011a, 2011b) include data obtained from the National Student Clearinghouse on college enrollments outside the public university system. If the rate of enrollment in such institutions is equal across participants and non-participants, conditional on other covariates, then the estimated effects of dual enrollment participation will tend to be attenuated toward zero (Hausman, Abrevaya, and Scott-Morton, 1998; Meyer and Mittag, 2013). The degree of attenuation will be approximately equal to the percent of students attending non-covered colleges, which we estimate at 17% in our dataset and is 10-15% in states considered in several existing studies (Speroni, 2011b; Struhl and Vargas, 2012).

Empirically, however, we observe that Running Start students are more likely to attend in-state public colleges than other similar students. This is not surprising given that they have already begun enrollment in a public college and the state public universities have transfer agreements with community colleges that allow transferred credits to satisfy specific core curriculum and major requirements (Washington Council for High School-College Relations, 2014). Thus, it is likely that, without data on enrollments in private colleges or out-of-state public colleges, estimates of dual enrollment effects are biased upward.

While our measure of college enrollment may be more complete than those used in previous studies, the NSC data has some well-known limitations (Dynarski et al., 2013).[14] We therefore begin by assessing the remaining measurement error in our data. In order to estimate the misclassification bias, we first need estimates of the probability of misclassification and the expected values of the covariates given misclassification (Meyer and Mittag, 2013). We describe the estimation procedure more fully in Appendix A, but briefly describe the assumptions we make. There are four sources of missing data in our analytical dataset. First, students who do not graduate high school are typically not submitted to NSC for matching. Nonetheless, some students who do not graduate high school do attend postsecondary institutions, usually at community colleges. We assume that students who do not complete high school do not attend colleges outside the Washington public college system. Second, approximately 3.5% of high school graduates were not submitted to NSC for matching. We assume that these students attend college outside the Washington public college system at similar rates as submitted students. Third, students may block the release of their records by NSC. Fourth, not all postsecondary institutions are covered by NSC. For these last two limitations, we assume that blocking and institutional coverage are independent of inclusion in our dataset conditional on attendance at a private or out-of-state college.

Using these assumptions, we estimate that a small percentage of students are misclassified as not attending college. Across all observations, we estimate that 2% are classified as not attending college but actually enroll, while we estimate that 1.7% are not properly classified as attending college full-time or attending a four-year institution. Following the method of Meyer and Mittag (2013), we estimate that these omissions cause an upward bias in all of our estimates. In **Table 3**, we display the bias estimates for each of our college enrollment outcomes. We estimate an upward bias of 1.1 percentage points for the college enrollment outcome, 0.8 percentage points for the full-time college enrollment outcome, and 0.6 percentage points for the full-time four-year college enrollment outcome. While the first two estimates are large relative to the coefficient estimates, they do not substantially change the college enrollment results.

We now compare our results to those of previous studies by omitting in-state private and out-of-state college enrollments from our college attendance measures. In **Table 3**, we repeat the baseline regressions in **Table 2** using only enrollments in state

[14] In particular, the NSC does not obtain records from all post-secondary institutions and some students block the release of their attendance records. In addition to these standard limitations, we also lack enrollment records for those students who fail to complete a high school degree, a small number of whom do enroll in postsecondary institutions. The effect of this latter category is likely to be small, particularly since they seem to enroll predominantly in public two-year institutions, which have open enrollment policies.

public two- and four-year colleges as our outcomes. Although the majority of college students are correctly classified, the omission of private and out-of-state enrollments substantially alters our estimates of college enrollment effects. Excluding private and out-of-state enrollments, the estimated effect of Running Start on any college enrollment increases from 0.053 in column (1) to 0.145 in column (2). We see similar increases in full-time and four-year enrollments, which increase from 0.004 to 0.080 and from -0.091 to -0.057, respectively.

In our sample, 72% of first-time college students enroll in a Washington State public college. It should be noted that the coverage rate for in-state publics is somewhat smaller than those reported in other studies, which typically have coverage of about 80% of college enrollments. We therefore provide estimates of the relationship between the misclassification bias and the probability of misreporting that supports the bias estimates in Table 3. The external validity of this assessment relies on the assumption that the average characteristics of students with missing college enrollment outcomes are similar in our data to those used in other states. Given the differences in eligibility requirements between Running Start and dual enrollment programs in other states, it is unclear how these results will translate to different contexts. However, there is reason to think that estimates based on the Washington data are actually conservative estimates of the biases in other contexts. If dual enrollment students in other states disproportionately enroll in public, in-state colleges to enroll in the college in which they take courses or to take advantage of transfer agreements, as they appear to do in Washington, then in other states with more selective programs, there may be a greater discrepancy between the in-state public and private or out-of-state enrollment rates between participants and non-participants who are otherwise appear similar based on observables. Based on the Washington data, we find that the misclassification bias on any college enrollment is approximately 56% of the probability of misreporting. We therefore interpret our results as suggesting that up to half of the enrollment effects estimated in previous studies may be explained by measurement error.

Some previous studies have relied on samples of high school graduates with retrospective information on dual enrollment participation. The results in **Table 2** suggest that dual enrollment may influence the likelihood that students complete high school and thus appear in the data. In order to assess the influence of restricting the sample to high school graduates, we repeat our analyses using this sample. We report the results of this exercise in column (2) of **Table 3**, where the analysis sample to high school graduates. Despite the loss of some students who drop out of high school, the results are similar to those with the full sample. However, we do find some noticeable differences when we replace the academic achievement variables with measures taken at the time of graduation. In column (3), we replace tenth grade GPA with the final cumulative GPA for high school graduates. Participation in Running Start may improve students' preparation for upper level high school courses and thereby increase their cumulative GPA. On the other hand, the grading standards in college-level courses are likely stricter than high school courses and Running Start students may therefore graduate with lower GPAs than they would had they taken only high school courses. Under either of these scenarios, inclusion of a post-treatment academic measure may bias the estimated effects of Running Start. When we include final GPA, we find that estimated Running Start effects tend to be larger than with tenth grade GPA by about 2 percentage points. Cumulatively, we find that restricting the sample to high school graduates and using post-treatment measures of academic achievement raise the point estimates for the college enrollment outcomes by 2.5-3 percentage points.

Conclusion

We use a state longitudinal data system to evaluate the influence of a popular dual enrollment program on college attendance. The data used in this study improves on that used in recent studies of dual enrollment programs by including a more complete record of college enrollment and a sample of students taken before participation in the program.

We find that students who participate in dual enrollment are more likely to attend any college immediately after high school graduation, but are no more likely to attend college full-time and are less likely to attend a four-year university. Moreover, the availability of richer data turns out to be quite important as we show that the estimated effects of Washington's dual enrollment program vary greatly depending on the types of postsecondary student information available. This suggests earlier studies lacking information on students' attendance at private or out of state colleges may have substantially misstated the benefits of dual enrollment programs.

We also find important heterogeneity in the outcomes of Running Start students depending on their initial preparation for college. Consistent with prior studies, we find that low-achieving participants later enroll in college at higher rates than expected. However, we also find that such students are more likely to drop out of high school or otherwise fail to complete a secondary diploma in four years. When considering eligibility requirements for dual enrollment programs, policymakers ought to balance the benefits of increased college access against the costs of high school dropout (Manski, 1989). In particular, this suggests the need for more evidence on the long-term degree completion outcomes of dual enrollment participants with low academic preparation.

Given the likelihood of selection into Running Start based on unobserved determinants of college attendance, we are cautious about drawing causal inferences based on the present analyses. In particular, our control vector is limited to high school academic variables and a limited set of socioeconomic variables. Using students' stated college intent at the time of entrance into Running Start, we present some evidence that selection may explain part of our results. In particular, we find that the diversionary effect of Running Start is likely biased downward by differences in pre-treatment college plans. We also find some evidence that the diversionary effect is strongest among students who have nearly completed an associate degree and that the short-run effect on enrollment may not be predictive of longer-run effects on baccalaureate completion.

The increasing prevalence of state data warehouses that combine records from K-12 and postsecondary agencies with private and out-of-state college attendance should provide an opportunity to better study dual enrollment programs, including more quasi-experimental evidence of program impacts. The use of samples taken before program participation with more complete information on college enrollment expands the range of causal questions that can be posed. In particular, rigorous analysis of the effects of dual enrollment participation on overall educational attainment remains an important topic for future research.

References

Adelman, C. (1999). *Answers in the toolbox: Academic intensity, attendance patterns, and bachelor's degree attainment* (No. PLLI-199-8021). Washington, D.C.: National Institute on Postsecondary Education, Libraries, and Lifelong Learning.

Adelman, C. (2006). *The toolbox revisited: Paths to degree completion from high school through college*. Washington, D.C.: U.S. Department of Education.

Allen, D. (2010). *Dual enrollment: A comprehensive literature review and bibliography*. City University of New York, Office of Academic Affairs.

Allen, D., & Dadgar, M. (2012). Does dual enrollment increase students' success in college? Evidence from a quasi-experimental analysis of dual enrollment in New York City. *New Directions for Higher Education*, (158), 11–19.

Allensworth, E., Nomi, T., Montgomery, N., & Lee, V. E. (2009). College Preparatory Curriculum for All: Academic Consequences of Requiring Algebra and English I for Ninth Graders in Chicago. *Educational Evaluation and Policy Analysis, 31*(4), 367–391.

An, B. P. (2013a). The impact of dual enrollment on college degree attainment: Do low-SES students benefit? *Educational Evaluation and Policy Analysis*, *35*(1), 57–75.

An, B. P. (2013b). The influence of dual enrollment on academic performance and college readiness: Differences by socioeconomic status. *Research in Higher Education*, *54*(4), 407–432.

Angrist, J. D. (2001). Estimation of limited dependent variable models with dummy endogenous regressors: Simple strategies for empirical practice. *Journal of Business & Economic Statistics*, *19*(1), 2–16.

Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics*, *121*(1-2), 343–375.

Attewell, P., & Domina, T. (2008). Raising the Bar: Curricular Intensity and Academic Performance. *Educational Evaluation and Policy Analysis, 30*(1), 51–71.

Aughinbaugh, A. (2012). The effects of high school math curriculum on college attendance: Evidence from the NLSY97. *Economics of Education Review, 31*(6), 861–870.

Bailey, T., & Karp, M. M. (2003). *Promoting college access and success: A review of credit-based transition programs*. New York: Community College Research Center.

Berger, A., Turk-Bicakci, L., Garet, M., Knudson, J., & Hoshen, G. (2014). *Early College, Continued Success: Early College High School Initiative Impact Study*. Washington, D.C.: American Institutes for Research.

Cameron, S. V., & Heckman, J. J. (1993). The nonequivalence of high school equivalents. *Journal of Labor Economics*, *11*(1), 1–47.

Cellini, S. R. (2006). Smoothing the transition to college? The effect of Tech-Prep programs on educational attainment. *Economics of Education Review*, *25*(4), 394–411.

Comay, Y., Melnik, A., & Pollatschek, M. A. (1973). The option value of education and the optimal path for investment in human capital. *International Economic Review*, *14*(2), 421–435.

Dougherty, C., Mellor, L., & Jian, S. (2006). *The relationship between Advanced Placement and college graduation*. Austin, TX: National Center for Educational Accountability.

Doyle, W. R. (2009). The effect of community college enrollment on bachelor's degree completion. *Economics of Education Review*, *28*(2), 199–206.

Dynarski, S. (2003). Does aid matter? Measuring the effect of student aid on college attendance and completion. *The American Economic Review*, 279–288.

Education Research and Data Center. (2010). *Participation in Postsecondary Education: Washington State High School Graduates, 2008-09*. Olympia, Wa.: Education Research and Data Center.

Hausman, J. A., Abrevaya, J., & Scott-Morton, F. M. (1998). Misclassification of the dependent variable in a discrete-response setting. *Journal of Econometrics*, *87*(2), 239–269.

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.

Joensen, J. S., & Nielsen, H. S. (2009). Is there a causal effect of high school math on labor market outcomes? *Journal of Human Resources, 44*(1), 171–198.

Kane, T. J. (2007). Evaluating the impact of the DC tuition assistance grant program. *The Journal of Human Resources*, *42*(3), 555.

Karp, M. M., Calcagno, J. C., Hughes, K. L., Jeong, D. W., & Bailey, T. R. (2007). *The postsecondary achievement of participants in dual enrollment: An analysis of student outcomes in two states*. St. Paul, MN: National Research Center for Career and Technical Education.

Klopfenstein, K., & Thomas, M. K. (2006). The link between advanced placement experience and early college success. Unpublished manuscript.

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.

Light, A., & Strayer, W. (2000). Determinants of college completion: school quality or student ability? *Journal of Human Resources*, *35*(2), 299–332.

Light, A., & Strayer, W. (2004). Who receives the college wage premium? Assessing the labor market returns to degrees and college transfer patterns. *Journal of Human Resources*, *39*(3), 746–773.

Long, M. C., Conger, D., & Iatarola, P. (2012). Effects of High School Course-Taking on Secondary and Postsecondary Success. *American Educational Research Journal, 49(2), 285–322*.

Long, B. T., & Kurlaender, M. (2008). Do community colleges provide a viable pathway to a baccalaureate degree? *Educational Evaluation and Policy Analysis*, *31*(1), 30–53.

Manski, C. F. (1989). Schooling as experimentation: a reappraisal of the postsecondary dropout phenomenon. *Economics of Education Review*, *8*(4), 305–312.

Marken, Stephanie, Gray, L., & Lewis, L. (2013). *Dual enrollment programs and courses for high school students at postsecondary institutions: 2010-11* (No. NCES 2013-002). Washington, D.C.: National Center for Education Statistics, U.S. Department of Education.

Meyer, B., & Mittag, N. (2013). *Misclassification in binary choice models* (No. CES 13-27). Washington, D.C.: Center for Economic Studies, U.S. Census Bureau.

Morgan, R., & Klaric, J. (2007). *AP Students in College: An Analysis of Five-Year Academic Careers* (No. 2007-4). New York, NY: College Board.

Planty, M., Bozick, R., & Ingels, S. J. (2006). *Academic pathways, preparation, and*

performance (No. NCES 2007-316). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.
Rose, H., & Betts, J. R. (2004). The effect of high school courses on earnings. *Review of Economics and Statistics, 86*(2), 497–513.
Rouse, C. E. (1995). Democratization or diversion? The effect of community colleges on educational attainment. *Journal of Business & Economic Statistics*, *13*(2), 217–224.
Sandy, J., Gonzalez, A., & Hilmer, M. J. (2006). Alternative paths to college completion: Effect of attending a 2-year school on the probability of completing a 4-year degree. *Economics of Education Review*, *25*(5), 463–471.
Speroni, C. (2011a). *High school dual enrollment programs: Are we fast-tracking students too fast?* New York: National Center for Postsecondary Research.
Speroni, C. (2011b). *Determinants of students' success: The role of advanced placement and dual enrollment programs.* New York: National Center for Postsecondary Research.
Stange, K. M. (2012). An empirical investigation of the option value of college enrollment. *American Economic Journal: Applied Economics*, *4*(1), 49–84.
Stinebrickner, R., & Stinebrickner, T. (2012). Learning about academic ability and the college dropout decision. *Journal of Labor Economics*, *30*(4), 707–748.
Struhl, B., & Vargas, J. (2012). *Taking college courses in high school: A strategy for college readiness*. Jobs for the Future.
Thomas, N., Marken, S., Gray, L., Lewis, L., & Ralph, J. (2013). *Dual credit and exam-based courses in U.S. public high schools: 2010-11* (No. NCES 2013-001). Washington, D.C.: National Center for Education Statistics, U.S. Department of Education.
Washington Council for High School-College Relations. (2014). *The Intercollegiate Relations Commission Handbook*. Olympia, WA: Washington Council for High School-College Relations.
Washington State Board for Community and Technical Colleges. (2010). *Running Start Finance Study Report*. Olympia, Wa: Washington State Board for Community and Technical Colleges.
Washington State Board for Community and Technical Colleges. (2011). *Running Start: 2009-10 Annual Progress Report*. Olympia, Wa: Washington State Board for Community and Technical Colleges.
Washington State Board for Community and Technical Colleges. (2012). *Running Start: 2010-11 Annual Progress Report*. Olympia, Wa: Washington State Board for Community and Technical Colleges.
Washington State Board for Community and Technical Colleges. (2013). *The role of transfer in the attainment of baccalaureate degrees at Washington public bachelor's degree institutions, class of 2011* (No. 13-5). Olympia, WA: Washington State Board for Community and Technical Colleges.

Table 1. Summary statistics

	RS Students		Non-RS Students		All Students	
	Mean	SD	Mean	SD	Mean	SD
Student outcomes:						
HS Dropout	0.03	(0.16)	0.04	(0.19)	0.04	(0.18)
GED	0.01	(0.12)	0.02	(0.15)	0.02	(0.14)
HS Diploma	0.93	(0.26)	0.88	(0.32)	0.89	(0.31)
Earns associate degree	0.14	(0.35)	0.00	(0.00)	0.02	(0.15)
Any college enrollment	0.73	(0.44)	0.56	(0.50)	0.59	(0.49)
Any full-time college enrollment	0.59	(0.49)	0.48	(0.50)	0.49	(0.50)
Any college enrollment (inc. AA)	0.75	(0.43)	0.56	(0.50)	0.59	(0.49)
Any full-time college enrollment (inc. AA)	0.62	(0.48)	0.48	(0.50)	0.50	(0.50)
Full-time four-year college enrollment	0.36	(0.48)	0.31	(0.46)	0.32	(0.46)
Running Start participant, 11th grade	0.65	(0.48)	0.00	(0.00)	0.10	(0.30)
Running Start participant, 12th grade	0.85	(0.35)	0.00	(0.00)	0.13	(0.34)
Total Running Start credits attempted	47.44	(31.29)	0.00	(0.00)	7.35	(21.13)
College intent: AA + BA	0.24	(0.42)	0.00	(0.00)	0.04	(0.19)
College intent: BA only	0.41	(0.49)	0.00	(0.00)	0.06	(0.24)
College intent: AA only	0.08	(0.27)	0.00	(0.00)	0.01	(0.11)
College intent: no degree	0.27	(0.45)	0.00	(0.00)	0.04	(0.20)
Student controls:						
Student gender: Female	0.41	(0.49)	0.52	(0.50)	0.50	(0.50)
Student race: Asian	0.11	(0.31)	0.08	(0.28)	0.09	(0.28)
Student race: Black	0.03	(0.17)	0.04	(0.20)	0.04	(0.20)
Student race: Hispanic	0.05	(0.22)	0.11	(0.31)	0.10	(0.30)
10th grade GPA	3.23	(0.62)	2.80	(0.84)	2.86	(0.82)
10th grade math WASL	0.45	(0.80)	-0.08	(1.01)	0.00	(1.00)
10th grade reading WASL	0.41	(0.89)	-0.08	(1.00)	0.00	(1.00)
10th grade writing WASL	0.37	(0.71)	-0.07	(1.03)	0.00	(1.00)
FRL Status, 10th grade	0.13	(0.34)	0.22	(0.41)	0.20	(0.40)
Bilingual Status, 10th grade	0.01	(0.09)	0.03	(0.18)	0.03	(0.17)
Primary language is not English	0.09	(0.29)	0.12	(0.33)	0.12	(0.32)
Title I targeted assistance	0.05	(0.21)	0.09	(0.28)	0.08	(0.27)
Gifted education	0.04	(0.19)	0.03	(0.17)	0.03	(0.17)
Number of AP courses taken in grade 9-10	0.12	(0.36)	0.08	(0.30)	0.08	(0.31)
Migrant student	0.00	(0.06)	0.02	(0.13)	0.01	(0.12)
Unexcused absences	1.68	(5.89)	1.89	(6.40)	1.86	(6.33)
N	8,586		46,810		55,396	

Table 2: The effects of Running Start on high school completion and college enrollment

	High School Completion			College Enrollment			
	Dropout	GED	Diploma	Any College	Full-time College	Full-time Four-year	No dip., no college
Panel A: Average Running Start Effects							
Running Start	0.011***	0.004***	-0.023***	0.053***	0.004	-0.091***	0.011***
	(0.002)	(0.001)	(0.003)	(0.006)	(0.007)	(0.006)	(0.003)
N	55,396	55,396	55,396	55,396	55,396	55,396	55,396
Panel B: Running Start Effects by Academic Achievement							
Running Start, Q1 index	0.045**	0.042**	-0.048*	0.079***	0.029	-0.020*	0.015
	(0.021)	(0.016)	(0.026)	(0.025)	(0.023)	(0.012)	(0.024)
Running Start, Q2 index	0.022***	0.004	-0.058***	0.122***	0.063***	-0.015	0.032***
	(0.007)	(0.005)	(0.012)	(0.016)	(0.017)	(0.010)	(0.010)
Running Start, Q3 index	0.010***	0.002	-0.026***	0.067***	-0.008	-0.071***	0.015***
	(0.003)	(0.003)	(0.007)	(0.013)	(0.014)	(0.011)	(0.006)
Running Start, Q4 index	0.008***	0.002	-0.014***	0.022**	-0.040***	-0.141***	0.004
	(0.003)	(0.002)	(0.005)	(0.010)	(0.011)	(0.012)	(0.004)
Running Start, Q5 index	0.001	0.001	-0.005	0.032***	0.016*	-0.115***	0.001
	(0.001)	(0.001)	(0.003)	(0.008)	(0.010)	(0.013)	(0.003)
N	55,396	55,396	55,396	55,396	55,396	55,396	55,396
Panel C: Running Start Effects by Academic Intent							
Running Start, AA + BA	0.010***	0.003	-0.025***	0.083***	0.027**	-0.135***	0.010**
	(0.003)	(0.003)	(0.006)	(0.010)	(0.011)	(0.012)	(0.004)
Running Start, BA only	0.009***	0.004**	-0.020***	0.069***	0.022**	-0.061***	0.009**
	(0.002)	(0.002)	(0.004)	(0.008)	(0.010)	(0.008)	(0.004)
Running Start, AA only	0.027***	0.006	-0.048***	0.039**	-0.029	-0.187***	0.030***
	(0.008)	(0.005)	(0.013)	(0.018)	(0.020)	(0.017)	(0.010)
Running Start, no degree	0.009***	0.006**	-0.019***	0.010	-0.034***	-0.072***	0.010*
	(0.003)	(0.003)	(0.006)	(0.010)	(0.011)	(0.009)	(0.005)
N	55,396	55,396	55,396	55,396	55,396	55,396	55,396

Notes: All regressions include the following variables: indicators for student sex, race, FRL, migrant status, whether English is the primary language, and 10th grade participation in targeted assistance (Title I) programs, bilingual education, gifted education, special education services, number of 9th and 10th grade Advanced Placement classes, and 10th grade GPA, unexcused absences, and standardized test scores in math, language arts, and writing. Regressions additionally include school fixed effects. Academic index is a weighted average of GPA and test scores where the weights are determined by a regression of college enrollment on the academic achievement variables using the non-Running Start sample. The academic intent variables are described in the text. Standard errors clustered by school in parentheses.

$***\ p < 0.01$, $**\ p < 0.05$, $*\ p < 0.10$.

Table 3: Sensitivity of results to missing data and sample selection

Dependent variable	Baseline	Est. Bias	Other Data Scenarios (1)	(2)	(3)
Any college enrollment or AA deg.	0.053***	0.011	0.145***	0.059***	0.077***
	(0.006)		(0.007)	(0.006)	(0.007)
Full-time college enrollment or degree	0.004	0.008	0.080***	0.009	0.032***
	(0.007)		(0.007)	(0.007)	(0.007)
Full-time four year college enrollment	-0.091***	0.006	-0.057***	-0.090***	-0.063***
	(0.006)		(0.006)	(0.007)	(0.007)
In-state public enrollments only			Y	N	N
High school graduates sample			N	Y	Y
Final cumulative GPA			N	N	Y

Notes: The estimated bias is the bias estimated from remaining data limitations using the methods described in the text and in Appendix A. In-state public enrollments code students as attending college only if the college is a Washington State public college. High school graduates sample includes only those students who obtain a high school diploma. Final cumulative GPA denotes the observed high school GPA at time of graduation. All regressions include the following variables: indicators for student sex, race, FRL, migrant status, whether English is the primary language, and 10th grade participation in targeted assistance (Title I) programs, bilingual education, gifted education, special education services, number of 9th and 10th grade Advanced Placement classes, and 10th (or 12th) grade GPA, unexcused absences, and standardized test scores in math, language arts, and writing. Regressions additionally include school fixed effects. Standard errors clustered by school in parentheses.
$***\, p < 0.01$, $**\ p < 0.05$, $*\, p < 0.10$.

Appendix A. Classification Errors in the NSC Data

June 18, 2014

There are several sources of possible classification error in our NSC data, many of which are documented in great detail by Dynarski et al. (2013). In addition to these are a few that are particular to our dataset. In many regressions, we code students by their full-time enrollment status, which is missing for some students in the NSC data. We do not count these students as being enrolled full-time in the analyses in the text. However, a simple bounding exercise suggests that the maximal bias from this source of misclassification is likely to be very small. We therefore focus on the other sources of missing enrollments in our data.

The NSC data includes information on enrollment for students at covered institutions who do not block release of their enrollment records. We further lack information on a number of students whose records were not submitted by ERDC to NSC. Most of these are students who did not complete an on-time high school diploma, but some are coded in our data as having graduated. We consider each of these sources of missing data in turn.

Note that all of our classification errors are false negatives; that is, we incorrectly classify some students who attend college as not enrolling. We first introduce some notation to ease the exposition. Let $D = 1$ indicate that a student is in our analytic sample; $H = 1$ denote that a student graduates high school; $C = 1$ denote that a student attends college; and $W = 1$ denotes a student attends college in a public postsecondary institution in Washington State. Then the forms of misclassification are:

- M1. Students whose records are not submitted to NSC because they do not have a standard high school diploma, but who do attend some postsecondary institution ($H = 0, S = 0, C = 1, W = 0$).
- M2. Students who obtain a high school diploma, but are not submitted for matching to NSC ($H = 1, S = 0, C = 1, W = 0$).
- M3. Students who attend an institution and blocked release of NSC records ($B = 1, C = 1, W = 0$).
- M4. Students who attend an institution not covered by NSC ($U = 1, C = 1, W = 0$).

Note that these are mutually exclusive so that the total probability of misclassification is summative over these possibilities.

We first assume that students who fail to graduate from high school do not attend private or out-of-state institutions:

$$Pr(H = 0, C = 1, W = 0) = 0.$$

If we instead assume that high school non-completers enroll in private or out-of-state institutions at a similar rate as high school graduates, we can estimate

$$\begin{aligned}
&Pr(H = 0, C = 1, W = 0) \\
&= Pr(C = 1, W = 0|H = 0)Pr(H = 0) \\
&= Pr(C = 1, W = 0|H = 1)Pr(H = 0) \\
&= Pr(H = 1, C = 1, W = 0)\frac{Pr(H = 0)}{Pr(H = 1)} \\
&\approx 0.15 \times 0.12 = 0.018.
\end{aligned} \tag{1}$$

As an extreme upper bound, therefore, we estimate that only 1.8% of students are misclassified as not attending college due to the non-submission of high school non-completers. This should be quite conservative since it does not appear that many students who fail to complete high school attend private or out-of-state institutions. Of the 1,036 students lacking a high school credential that were submitted to NSC, only 17 (1.6%) attended a college covered by that sample. This suggests a misclassification probability of about 0.0016 (0.10×0.016), so we assume this probability is 0 for estimation purposes.

We begin by estimating the probability that a student has a record submitted to NSC and attends college outside the Washington public university system conditional on inclusion in our analytic dataset. This probability will be useful for several of the calculations later on. Let A_0 denote this event:

$$A_0 = \begin{cases} 1 & \text{if } C = 1,\ W = 0,\ S = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Note that this is complicated by the fact that we do not observe college attendance for all of these students; some students with NSC submissions attend uncovered institutions or have their records blocked from release to third parties. Nonetheless, we can use two pieces of available information to estimate this probability under reasonable assumptions. The first is the coverage rates estimated by Dynarski et al. (2013), which can be used to provide an estimate of the coverage rate conditional on attendance outside the schools covered by the ERDC data. The second is the NSC blocked records reports for the ERDC submissions, which can be used to estimate the blocked record rate conditional on attendance at a non-ERDC school. Putting these together, we have

$$\begin{aligned}
Pr(A_0 = 1|D = 1) &= Pr(A_0 = 1, B = 0, U = 0|D = 1) \\
&\quad + Pr(A_0 = 1, B = 1, U = 0|D = 1) \\
&\quad + Pr(A_0 = 1, U = 1|D = 1) \\
&= Pr(A_0, B = 0, U = 0|D = 1) \\
&\quad + Pr(B = 1, U = 0|A_0 = 1, D = 1) \times Pr(A_0 = 1|D = 1) \\
&\quad + Pr(U = 1|A_0 = 1, D = 1) \times Pr(A_0 = 1|D = 1).
\end{aligned}$$

Collecting terms and rearranging, we have

$$\begin{aligned}
Pr(A_0 = 1|D = 1) &= \frac{Pr(A_0 = 1, B = 0, U = 0|D = 1)}{1 - Pr(B = 1, U = 0|A_0 = 1, D = 1) - Pr(U = 1|A_0 = 1, D = 1)} \\
&\equiv \frac{A_1}{1 - A_2 - A_3}.
\end{aligned} \tag{2}$$

The first term, A_1, is the probability that a student is observed in a private or out-of-state college in our dataset and is therefore easily estimated from the sample. The other two probabilities are not directly observed but can be estimated from available data. The second, A_2, is the probability that a student in our dataset attends an out-of-state or private college at an institution covered by NSC and blocks release of her record. The third, A_3, is the probability that a student in our sample attends an institution not covered by NSC.

We begin by estimating the probability that a student attends an uncovered institution conditional on being in our dataset and attending a college outside the Washington public college system, $Pr(U = 1|A_0 = 1, D = 1)$. While we do not have any direct information about students' attendance at such institutions, we do have estimates of the NSC institutional coverage rates from Dynarski et al. (2013) for states and sectors and data from IPEDS about the enrollment of Washington State students in each state and sector. From these data sources, we construct an estimate of the probability that a Washington State student who enrolls in a college outside the public university system enrolls in an uncovered institution. To do so, we first use the 2011 IPEDS fall enrollment counts for recent high school graduates to estimate the probability that a student who enrolls outside Washington's public colleges enrolls in each state, level, and sector (public, private not-for-profit, private for-profit). We then estimate the total coverage rate for Washington students by multiplying each weight by the sectoral coverage rates reported in Tables A2 - A6 of Dynarski et al. (2013). Taking the sum over all sectors, we arrive at an estimate of $Pr(U = 1|C = 1, W = 0)$.

Under the assumption that $U \perp (D, S)|(C, W)$, we can use this as an estimate of the desired probability $Pr(U = 1|A_0 = 1, D = 1)$. That is, we assume that NSC coverage is independent of inclusion in our dataset and whether the student was submitted to NSC conditional on whether the student attends college outside the public system. This assumption is likely violated for two reasons.

First, the sample restrictions we impose on our data restrict the sample to students with above-average academic performance. We may therefore undercount students attending private, four-year colleges, which have a lower coverage rate. On the other hand, the IPEDS data include students who attend private high schools in Washington, which should tend to bias $Pr(U = 1|C = 1, W = 0)$ upward if such students are more likely than public high school students to attend private colleges. Note that these biases tend to operate in the opposite direction. Nonetheless, our estimated non-coverage rate is very close to both the Washington and national averages, and using these rates instead has little impact on the results.

It remains to estimate the probability of FERPA blocking. The NSC returns a report that indicates the number of blocked records. Because we only know the institution and blocked record rate from this data, we only have an estimate of

$$Pr(B = 1|A_0 = 1, U = 0),$$

whereas the bias term requires an estimate of $Pr(B = 1, U = 0|A_0 = 1, D = 1)$. We therefore estimate

$$\begin{aligned} &Pr(B = 1, U = 0|A_0 = 1, D = 1) \\ &= Pr(B = 1|U = 0, A_0 = 1, D = 1)Pr(U = 0|A_0 = 1, D = 1) \\ &= Pr(B = 1|U = 0, A_0 = 1)Pr(U = 0|A_0 = 1), \end{aligned} \tag{3}$$

where the second inequality follows from the same assumptions as above plus the assumption that $B \perp D|(A_0, U)$. With these estimates in hand, we can estimate the probability of misreporting due to M3 and M4.

The last source of measurement error is students who complete high school but are not submitted to NSC. We estimate the probability of misclassification under the assumption that the probability of college enrollment is independent of whether the student is submitted to NSC conditional on high school graduation. Note that this uses the previous assumption that $Pr(C = 1, W = 0|H = 0) = 0$. We then have

$$\begin{aligned} &Pr(C = 1, W = 0, S = 0|D = 1) \\ &= Pr(C = 1, W = 0, S = 0, H = 1|D = 1) \\ &= Pr(C = 1, W = 0|H = 1, S = 0, D = 1)Pr(H = 1, S = 0|D = 1) \\ &= Pr(C = 1, W = 0|H = 1, S = 1, D = 1)Pr(H = 1, S = 0|D = 1) \\ &= Pr(C = 1, W = 0, H = 1, S = 1|D = 1)\frac{Pr(H = 1, S = 0|D = 1)}{Pr(H = 1, S = 1|D = 1)} \\ &= Pr(A_0 = 1|D = 1)\frac{Pr(H = 1, S = 0|D = 1)}{Pr(H = 1, S = 1|D = 1)} \\ &\approx 0.0053. \end{aligned} \tag{4}$$

Note that the first term is the same we estimated earlier. The other terms can be estimated directly from the sample.

Finally, we combine these to estimate the total misreporting rate. We have

$$\begin{aligned}
Pr(M=1|D=1) &= Pr(M1=1|D=1) + Pr(M2=1|D=1) \\
&+ Pr(M3=1|D=1) + Pr(M4=1|D=1) \\
&= 0 + Pr(A_0=1|D=1)\frac{Pr(H=1,S=0|D=1)}{Pr(H=1,S=1|D=1)} \\
&+ [Pr(B=1,U=0|A_0=1,D=1) + Pr(U=1|A_0=1,D=1)] \\
&\times \frac{Pr(A_0=1,B=0,U=0|D=1)}{1-Pr(B=1,U=0|A_0=1,D=1)-Pr(U=1|A_0=1,D=1)}.
\end{aligned} \tag{5}$$

Using this data, we can attempt to recreate the biases in the linear probability model using the formulas of Meyer and Mittag (2013). We first summarize the assumptions we have made thus far.

A1. $Pr(C=1,W=0,H=0)=0$.

A2. $U \perp (D,S)|(C,W)$.

A3. $B \perp (D,S)|(C,W)$.

A4. $(C,W) \perp S|(D,H)$.

Following Meyer and Mittag (2013), define the measurement error in the dependent variable as

$$u_i = y_i - y_i^T = \begin{cases} -1 & \text{if } i \text{ is a false negative} \\ 0 & \text{if } i \text{ is reported correctly} \\ 1 & \text{if } i \text{ is a false positive} \end{cases}$$

The bias in the OLS coefficients is given by

$$\delta = (X'X)^{-1}X'u,$$

which, given the structure of the measurement error, can be rewritten as

$$\delta = -(X'X)^{-1}N_{FN}\bar{x}_{FN},$$

where N_{FN} is the number of false negatives and $\bar{x}_{FN}$ is the means of the covariates among the false negatives. Our estimate of the bias then becomes

$$\hat{\delta} = -N(X'X)^{-1}Pr(M=1|D=1)\bar{x}_{FN},$$

where the mean is taken over the observations that are observed to attend college outside the Washington public system ($C=1,W=0,S=1,B=0,U=0,D=1$).

References

Dynarski, S. M., Hemelt, S. W., and Hyman, J. M. (2013). The missing manual: Using National Student Clearinghouse Data to track postsecondary outcomes. Technical Report 19552, National Bureau of Economic Research, Cambridge, MA.

Meyer, B. and Mittag, N. (2013). Misclassification in binary choice models. Technical Report CES 13-27, Center for Economic Studies, U.S. Census Bureau, Washington, D.C.